科学结构的表征与不变性

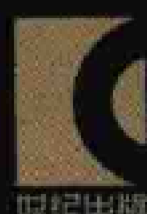
〔美〕帕特里克·苏佩斯 著　成素梅 译

二十世纪西方哲学经典

Patrick Colonel Suppes

Representation and Invariance of Scientific Structures

上海译文出版社

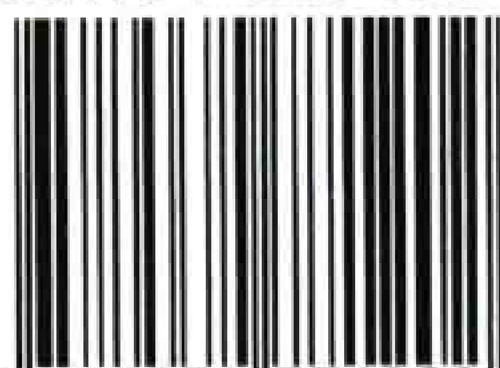


《二十世纪西方哲学经典》选收二十世纪西方哲学界各主要流派影响较大的著作，通过有选择的译介，旨在增进文化积累，拓展学术视野，丰富研究课题，为了解和研讨现代西方哲学提供系统而完整的第一手资料，以利于理论界、学术界深化对西方文化的研究和借鉴。

在科学哲学中运用形式方法的一个根本原因是拥有一个固定的指称框架的可取性，这种指称框架可以被用来组织手头的许多学说。《科学结构的表征与不变性》考察了集合论方法是如何提供这样一种指称框架的，全书内容涵盖了公理化方法、表征、不变性、概率、力学以及语言等问题，同时还包括对关于语词和句子的脑波表征的研究。本书是帕特里克·苏佩斯教授的代表作，是其数十年研究成果的结晶，是了解科学哲学当前进展的必读作品。

上架建议：外国哲学
ISBN 978-7-5327-7120-2





关注下载 译文APP
名家名著 一手掌握
上海译文出版社
www.yiwen.com.cn
定价：108.00元
易文网：www.ewen.co

科学结构的表征与不变性

〔美〕帕特里克·苏佩斯 著　成素梅 译

二十世纪西方哲学经典

Patrick Colonel Suppes

Representation and Invariance of Scientific Structures

上海译文出版社

图书在版编目(CIP)数据

科学结构的表征与不变性/(美)苏佩斯
(Suppes, P. C.)著;成素梅译. —上海:上海译文出
版社, 2016.1
(二十世纪西方哲学经典)
ISBN 978-7-5327-7120-2

Ⅰ.①科… Ⅱ.①苏… ②成… Ⅲ.①科学体系学-
研究 Ⅳ.①G304

中国版本图书馆 CIP 数据核字(2015)第 271133 号

Patrick Colonel Suppes
Representation and Invariance of Scientific Structures

本书根据美国斯坦福大学语言与信息研究中心 2002 年版译出

图字:09-2009-564 号

科学结构的表征与不变性
[美]帕特里克·苏佩斯 著 成素梅 译
责任编辑/王巧贞 封面设计/张志全工作室

上海世纪出版股份有限公司
译文出版社出版
网址:www.yiwen.com.cn
上海世纪出版股份有限公司发行中心发行
200001 上海福建中路 193 号 www.ewen.co
山东鸿杰印务集团有限公司印刷

开本 890×1240 1/32 印张 26.25 插页 5 字数 533,000
2016 年 1 月第 1 版 2016 年 1 月第 1 次印刷
印数:0,001—1,000 册

ISBN 978-7-5327-7120-2/B·419
定价:108.00 元

译者前言*

迅猛发展了半个多世纪的科学哲学研究进路，经过科学哲学自身发展逻辑的内在演变，经过科学的人文社会学研究进路的外在冲击，目前正处于转型之中。转型中的科学哲学一方面揭示了标准科学哲学研究进路的局限性，另一方面正在以一种开放的兼收并蓄的态度探索新的研究进路。如果说我们把20世纪科学哲学面临的矛盾大致概括为绝对主义与相对主义的对立、科学主义与人文主义的对立、实在论与反实在论的对立、规范的科学哲学与描述的科学哲学的对立，等等，那么到21世纪，如何真正超越这种非此即彼的二值逻辑的思维模式，便成为内在论者的科学哲学研究之重点，或者说，如何寻找一个新的视角来重新理解科学，成为当代科学哲学发展的重要方向之一。本书为思考这一问题提供了一个可供参考的视角。也许正因为如此，本书出版不到一年就荣获了2003年度的科学哲学拉卡托斯奖。

一

本书的作者帕特里克·苏佩斯(Patrick Suppes, 1922—)是美国逻辑学家、科学哲学家和心理学家，在逻辑学、科学哲学、

---

* 本前言曾作为书评发表于《哲学分析》杂志创刊号(即，2010年第1期)。

测量理论、量子力学基础、决策理论、心理学和教育技术等领域做出了重大贡献。他最初毕业于芝加哥大学,学习气象学,后来在哥伦比亚大学师从内格尔(Ernest Nagel)学习哲学,并于1950年获得博士学位,1952年到斯坦福大学执教至今。1959年到1992年担任斯坦福大学社会科学数学研究所(IMSSS)所长职务,现在是斯坦福露西·斯特恩(Lucie Stern)哲学教授。1965年由于他在数学心理学方面的工作被推选为国家科学院院士,1972年到1973年担任美国哲学家协会主席,1972年获得美国心理学协会杰出科学贡献奖,1973年到1974年担任美国教育研究协会主席,1973年到1977年担任国家教育科学院院长,1975年到1979年担任科学史与科学哲学国际协会逻辑、方法论与科学哲学专业委员会主席,1976年和1978年担任国际科学史与科学哲学协会主席。1990年获得由布什总统颁发的权威性的国家科学家勋章,1994年成为计算机协会成员。现在年近90高龄的他仍然领导着一个研究小组通过语言的脑电图表征实验研究认知科学问题。这个小组的成员具有互补的知识结构(比如,物理学、生物学、生物医学等背景),来自美国、英国、法国、中国以及葡萄牙等国家。

本书长达808页,是作者的一本精心之作。正如他在前言中讲到的,“我着手写作本书时,我还是一位年轻人,对,作为年轻人,当然我至少还不到四十岁。本书完成于我退休十年后的八十岁。”①从这一段话中也可以看出,作者对本书的构思开始于科学哲学的逻辑经验主义学派向历史主义学派的转型时期。这也决定了本书论证问题的出发点必然是以批判逻辑经验主义

① Patrick Suppes, *Representation and Invariance of Scientific Structures*, California: CSLI Publications, 2002. p. xv.

为前提。所不同的是,他对多元论与整体论观点的阐述,没有采取单纯的历史主义进路,而是以历史追溯为背景、以最新实验为手段的集合论进路。因此,本书对问题的论证方式,既不同于只以科学史的文本资料考证为基础的大多数科学哲学著作,也不同于单纯以跟踪描述科学实验活动为核心的科学知识社会学著作,而是立足于集合论进路,运用公理化方法,把哲学思考、文本考据和最新的认知神经科学实验有机地结合起来,跨学科、多视域地论证科学哲学命题。其论证范围涉及哲学、集合论、逻辑学、概率论、几何学、物理学、心理学、认知科学、神经科学、人工智能等领域,富含许多深邃的科学哲学见解,既是作者多年研究成果的汇集,也是一本思想性极强的科学哲学巨著。

本书贯穿始终的一个基本思想是,试图用集合论的进路讨论一般的科学哲学问题。作者认为,集合论进路的重要优势有二。其一,为研究过去或当前科学的任何成体系的表征和不变性问题提供正确的方法,从而避免运用人工语言讨论科学哲学问题所带来的不足;其二,能够把在一般论述层面上运用传统方法很容易丢失的差异突出出来,试图在一种更具有可操作性的实验层面讨论过去认为是形而上学的哲学观点。本书突出运用的基本研究方法是,以集合论进路为基础的公理化方法;以概念追溯为线索的历史分析方法,以及用脑电图表示词和句子的实验为根据的实验研究与数理统计方法。本书围绕对理论结构、主观概率与客观概率、还原论、模型的同构、物理学中的不变性、视觉空间的不变性、强弱可逆性、心理表征、机器学习、刺激-反应模型等概念的历史性阐述,重点回答了“为什么表征或不变性在科学基础或科学哲学中是重要的?”这个敏感问题。一个重要的特点是,以举例说明的方式来论证抽象的问题。

就像一部好的影片能让不同兴趣的观众获得乐趣,一部好

的小说能让不同水平的读者产生共鸣一样，一本有学术价值的科学哲学著作也应能吸引不同领域的读者。这本书就具备了这样的特点。首先，数学哲学家可以从中体会到如何基于集合论进路和运用公理化方法来论证哲学问题，以及从关于概率论和公理化方法的历史发展、几何作图的视错觉效应（例如，著名的凸肚状的建筑设计）、几何学中的有穷论、几何的层次结构以及视觉空间是不是欧几里得空间等问题的讨论中获得启迪；其次，物理哲学家可以从中体会到如何根据几何的集合论进路并运用群论方法阐述时空问题和物理学中的不变性与协变性问题，隐变量的量子理论以及从关于对称性、不变性、还原论、因果性、随机性和可逆性等概念的例证讨论中得到启发；第三，心理学家可以从中看到如何通过理论联系实验的方式论述刺激-反应理论的发展及其与自动机理论之间的关系；第四，语言学家可以从中看到如何通过心理学中的联想或条件作用的简单机制来充分地说明复杂而微妙的语言（主要指计算机程序语言）现象，以及机器学习等问题；第五，也是最重要的一个方面是，科学哲学家能够从中看到如何基于对传统的科学理论观的评判，运用公理化方法，立足于对不变性与表征概念的阐述，论证一种多元论的理论观。

二

“科学理论是什么”的问题，是科学哲学的一个基本问题，也是本书讨论问题的出发点。在科学哲学的发展史上，逻辑经验主义提供的科学理论观被誉为“标准的理论观”。这种理论观通常把科学理论看成是由两部分构成的，一部分是抽象的逻辑计算，包括逻辑学的词汇和理论的原始符号在内。理论的逻辑结构是根据理论的原始符号、陈述理论的公理或公设来确定的。

对于许多理论而言,原始符号被认为是像"电子"或者"粒子"之类的理论术语,这些术语不可能以任何简单的方式与可观察的现象直接相关,需要有卡尔纳普(Rudolf Carnap)所说的对应规则。因此,理论的第二部分是为理论术语提供"操作定义"或"经验解释",即赋予逻辑计算以经验内容。这种科学理论的标准纲领是从近代物理学的研究方式中抽象出来的。[①] 从科学史与当代自然科学的研究方式来看,这种观点虽然不是完全错误的,但却过分简单,既遗漏了理论的某些重要特性,也遗漏了理论之间存在的内在差别,因而从一开始就受到了许多批评。

本书在绪论部分基于对这种传统理论观的批判,提出了把科学理论看成是一个模型集合的观点。这也是作者多年来一直提倡的基本进路。他认为,传统的理论观在逻辑的意义上不可能为形式计算提供一种适当的语义学,或者说一个复杂理论的句法结构通常很少提供关于理论本性的理解。相比之下,把理论看成是一个模型集合,并用理论模型的术语分析理论结构,除了容易避免传统理论观所遇到的难题之外,还能把较强的数学元素引入讨论中,使我们有可能在一阶逻辑的范围内阐述理论,即通过定义集合论谓词的方法使一个科学理论公理化。在公理化的基础上,以许多不变性事例为基础,讨论理论模型之间的同构问题和理论的还原问题,进而达到为一个科学理论证明一个表征定理的目的。这条进路的意图在于通过理论模型之间存在的同构关系证明理论的表征定理,通过理论表征定理的证明在统计学意义上为科学实在论辩护。

对科学实在论观点的这种辩护方式比科学实在论的"无奇

① 参见,成素梅,"逻辑经验主义的理论观及其影响",《社会科学》2009年第1期。

迹"论证和"操作论证"更具有说服力。因为它既不是以科学的成功事例为背景进行单纯的归纳劝导,也不是以观察实验操作为基础进行人文社会学的描述,而是根据当代心理学与认知科学的最新进展,对认知神经科学实验结果进行统计学分析得出的结论。这就使得我们对科学的实在论辩护有可能超越观念层面的争论,具有了实证的意义,也为近代以来贝克莱与休谟等哲学家关于"一般思想是否存在"的哲学争论,提供了一种有说服力的解答。这是本书第八章("语言的表征")重点论述的内容。作者在这一章占用了很长的篇幅介绍了他领导的实验小组在20 世纪 90 年代末的实验结果。他们的实验表明,在计算机屏幕上产生的熟悉形状的视觉图像,比如,一个圆或三角形,非常类似于由相应的口语名词产生的大脑图像,从而在统计学的意义上为爱因斯坦所说的纯粹概念能够把握实在的哲学观点提供了可靠依据。

这种运用现代科学手段间接地论证传统哲学结论的研究方式,与正在兴起的科学实践哲学的转向相呼应。但侧重点完全不同。当代的科学实践哲学家虽然大多数人具有自然科学背景,但他们在通常情况下并不亲自从事科学实验研究,更没有自己的实验室,只是通过对科学史案例的重新剖析或通过在他人的科学实验室的参与性观察,对什么是科学实践、如何建构科学实践以及科学实践成功的条件等一系列问题给予新的哲学关注。他们试图基于分析的框架考虑理论,基于实践的基础考虑世界,或者说,他们试图为传统的科学哲学补充"实践"维度,为科学的人文社会学研究补充"世界"维度,从而达到对理论、实践与世界三者关系的更合理的理解。从这个意义上看,本书采取的集合论进路已经起到了异曲同工之作用。

作者认为,回答科学理论是什么的问题并不是给出这样的

答案,"X是一个科学理论,当且仅当,某某"。而重要的是承认,存在着由检验基本理论的实验方法所产生的理论的层次结构。理论的这种层次结构体现了理论的复杂程度。把科学理论看成是一个模型集合的优势之一就是能够根据理论的复杂程度对理论进行分层,不同层面的理论对真理概念具有不同的要求与理解。这是一种多元论的理论观。在传统的科学理论观中,关于真理概念的讨论非常单一,也没有可操作性。实际上,在科学实践活动中,理论的重要功能不只像传统理论观所阐述的那样,是组织陈述或断言陈述的真伪。从实用主义的角度来看,不同层次的理论具有不同的功能与作用,作者举例说明了复杂度不同的理论具有的三种作用:其一,把理论理解为一组推理原理;其二,把理论理解为组织证据的方法;其三,把理论理解为依赖于语境的一组模型。

例如,在熟悉的三段论的推理中,"所有的人都是会死的;苏格拉底是人;因此,苏格拉底是会死的"。按照工具论的观点,大前提"所有的人都是会死的",被转化为一个推理原理。当把理论看成是推理原理,而不是大前提时,虽然我们不再直接关心确定理论的真伪,而是评价理论在推断新的事实陈述时的有用性,但这并不是最重要的。重要的是,如何为替代经典的真理概念提供一个适当的语义学。为此,作者进一步考察了从关注理论陈述转向关注行动的统计决策理论的情况。他认为,一旦关注问题的视角从陈述转向行动,就自然地用预期损失或风险概念取代了真理概念。问一个陈述是否正确,是适当的,但是,问一个陈述是否有风险,就没有意义。相比之下,问一种行动有多大风险,而不是问这种行动是否正确,是合理的。在这种情况下,理论成为组织证据和选择行动的方法。但这种观点同样没有提供一种新的理论观,也不能用来评价更复杂的科学理论。因为

这种工具论的理论观把理论与事实区分开来,只关注语义问题,而忽略了更重要的语用问题。

最后,作者站在心理学的行为主义的立场上,认为现代逻辑的语义分析,由于没有明确地考虑说话者、作者、听众和读者对语言刺激的产生与接受,因而不足以对语言的认知用法给出说明,更无法对"什么是科学理论"这个问题提供最终有意义的答案。一个适当而完备的回答,只能根据理论的产生者和使用者的明确而详细的考虑来提供。[①] 如果从行为主义的立场上来看待理论,就有可能把模型和理论作为一级近似来谈论,从而把理论理解为一个模型集合。作者把他的这种理论观称为是一种更根本的工具性的理论观,但并没有停留于此。作者在第八章从行为主义的心理学分析转向把语言分析延伸到具有可操作性的脑电图表征的认知神经科学的分析。特别是通过对机器学习与心理学中的刺激-反应模型的比较讨论,在新的层面上回到了对关于真理问题的实证分析层面,从而有可能在一个更高的层面把科学哲学中的实用主义、工具主义、操作主义、实证主义和实在论等观点有机地统一起来。

三

本书对科学结构问题的探讨,是通过对表征与不变性概念的详细阐述进行的。表征概念在日常生活中是常见的。从理论上讲,对某物的表征是指一种映像、模型或对它的复制。表征的形式理论或数学理论的主要目标是深化我们对问题的理解。表征与还原概念相关。在哲学史上,把一类现象或一组想法还原

---

① Patrick Suppes, *Representation and Invariance of Scientific Structures*, California: CSLI Publications, 2002. p. 9.

为另一类现象或另一组想法的哲学主张，是一个很古老的话题。伊壁鸠鲁就曾提出了把宇宙中的复杂物体还原为简单物体的主张，笛卡儿阐述了几何还原为代数的观点，物理学中典型的事例是把热力学还原为统计力学，隐变量的量子理论也是以还原论的思想为基础提出的，近代以来的许多哲学家甚至主张把心理学还原为生理学，把生理学还原为物理学，逻辑经验主义主张把理论陈述还原为观察陈述，而某些科学的人文社会学研究者则反过来主张把物理学还原为社会科学，等等。在作者看来，所有这些还原的观点，事实上都只是某种推测。作者证明，在可计算的数学意义上，即使是非常简单的还原，也是难以实现的。这也是他自己最终没有在本书中为把一个理论还原为另一个理论的观点提供有力支持的原因所在。然而，尽管如此，作者并没有因此而彻底否定还原的可能与价值，而是采取了一种多元论的宽容态度，为还原的可能留有余地。

除了还原概念之外，与表征思想复杂地联系在一起的是表征的不变性思想，而不变性概念与对称性概念有着密切的关系，例如，一个正方形围中心旋转 90 度，形状不变，是因为正方形具有对称性，而一个三角形只有在旋转 180 度的条件下才是不变的。对称性思想又可以用数学中的群概念来描述。日常生活中有关对称性的事例是极其丰富的，比如，我们人体是对称的，许多古典建筑与美术设计图案是对称的，对称性也是许多物理学定律提出的一个基本前提，等等。作者用数学术语把这种对称性表述为："一个几何图形在给定群的条件下是**对称的**，如果这个群的每个自同构都把这个图形映射到它自身的话。"[1]这里的

---

① Patrick Suppes, *Representation and Invariance of Scientific Structures*, California: CSLI Publications, 2002. p. 100.

映像概念蕴含了一种变换关系,在一种变换关系下的不变性,是证明一个理论存在表征定理的基本前提。不论是在测量理论中,还是在物理学中,不变性都是很重要的概念。本书第四章以丰富的事例说明了这一连串相关概念的相互依赖关系。

接下来的第五章到第八章着重讨论了表征在科学哲学中四个重要的思想领域的具体应用,它们是:概率的表征、时间与空间的表征、力学的表征以及语言的表征。在这几章里,作者对许多概念的阐述很独特,例如,第五章在讨论主观概率与客观概率问题时,详尽地论证了倾向性不是概率的观点和未知概率的问题;第六章从物理学转向心理学,最后,再转向几何,其中,关于视觉现象的讨论是很有新意的;第七章在讨论力学的表征问题时,以牛顿力学中的受限制的三体运动为例,证明了决定论的力学系统也会出现随机解,并举例说明了因果性过程的强可逆性与弱可逆性概念;第八章阐述了用联想或条件作用的简单机制说明复杂而微妙的语言现象的可能性,等等。

总之,科学理论的表征和不变性既是科学哲学中非常重要的问题,也是当代认知科学研究的核心问题。本书从跨学科的视域,运用逻辑、集合论和实验统计等方法,把科学哲学的研究视野与当前的自然科学的最近发展真正地联系起来,把多元论的视角与实用主义的视角联系起来,把理论表征与模型建构联系起来,系统地提供了一条科学哲学研究的集合论进路。本书对于从事科学哲学、心智哲学、认知科学、数学心理学和统计学基础理论以及心理学研究的学者与爱好者都有重要的参考价值,是一本思想性强和资料翔实丰富的跨学科视域的科学哲学著作。

中文版序言

很荣幸出版我的书的中文版。近几年来，我个人招生了好多很有能力的中国学生。我愿意认为，我的这本著作对于更广泛范围的这些学生来说是有益的。

原英文版的前四章一般地介绍了科学中的表征概念和不变性概念，我希望对中国学生来说这同样也是有用的。后四章是一般观点的应用，它们也可能会引起这样一些学生的兴趣：这些学生受过良好的科学教育，但不太熟悉关于特别是物理学和心理学领域的主要问题的基本讨论。

本书的一个重要特征是强调各种特殊科学理论的经验细节和理论细节。我的希望是，中国读者将不只是在研究一般性问题时，而且还在研究科学方法论的特殊细节（比如，测量理论）或基本论题（比如，量子力学中的隐变量的存在和纠缠的本性）时，分享这种兴趣。我试图以非专家的普通读者容易理解的方式来写这些问题。我也试图以各种方式把历史背景带入许多基本科学概念的讨论中。

最重要的是，我要感谢我的译者成素梅教授，为了提供准确的和容易理解的中文版，她付出了长期努力。

帕特里克·苏佩斯

2010 年 11 月

美国斯坦福

前　言

本书写了很长时间。我最早的预备版本可追溯到 1962 年临时装订成册的标题为《科学中的集合论结构》(*Set-theoretical Structures in Science*)一书,但我知道,更早的草稿源于 20 世纪 50 年代我在斯坦福大学给本科生开设的科学哲学课。课程讲义在概念上沿着我于 1957 年首次出版的《逻辑导论》(*Introduction to Logic*)的最后一章关于"公理化方法的集合论基础"的思路很快展开。 xiii

我还记得,在那些早期岁月里,在几种场合下都有人问我,使集合论结构在科学哲学中起核心作用的一般观点是什么呢?最初,我强调地回答说,在欧几里得的《几何原本》(*Elements*)出版很久以前,公理化方法的许多一般的智力优势就很明显。但我逐渐明白,还有一种更具有哲学意味的更好的回答。那就是,这样的结构为研究过去或当前科学的任何成体系的部分中的表征(representation)和不变性(invariance)问题提供了合适的背景。当然,这种回答带来了另外一个问题:表征或不变性为什么在科学基础或科学哲学中是重要的呢?从某种意义上说,这正是整个这本书要回答的问题。但一些标准的事例可能是有帮助的,即使很不详细。

19 世纪伟大的智力胜利之一是,只根据粒子运动的表征对

像温度和压力之类的熟悉概念作出力学说明。带有更多困扰的一个同样伟大的胜利是,在20世纪初实现了必须用爱因斯坦的狭义相对论的时空及其新的不变性来取代经典物理学中标准的即使通常只是被默认的空间与时间相分离的不变性。

但这些事例不只来源于物理学。柏拉图和亚里士多德的著作中有对知觉中的表征本性的精致分析,而且这种分析在当代心理学和哲学的争论中仍然很盛行。

因此,我不会为我在修订本书的题目时反映出对表征和不变性的强调而感到歉意。在不同类型的基础研究中,古往今来的核心论题都很容易归入这个标题之下。当然,这无疑不是科学哲学的全部,而只是一个主要部分。

xiv

前四章一般地介绍了表征和不变性概念。后四章提供了在科学哲学中重要的四个思想领域的应用,而说重要,是因为这些领域具有更普遍的科学意义。它们是概率的本性、空间与时间概念、在经典力学和量子力学中的物理表征和心理表征,最后,从几个不同的视角考察了语言的表征。我在本书的最后总结出不同章节中讨论和陈述的表征和不变性定理的一个汇总表。许多定理没有证明,只描写了数学与科学文献中的著名结果。通常(但不总是)证明了的那些定理代表了我自己工作的某些方面,而且从数学的立场来看,大多数给出的证明是初步的。因为任何一位执著的读者显然主要关注的是概念分析与分类,而不是形式证明。

尽管我一直致力于科学哲学中的形式方法,但是还存在着我认为几乎是可靠的另外两条进路,所以这两条进路也一直很有影响。一是关注经验细节。我从不同方面考虑的许多实验,

特别是心理学实验,反映了这一点。此外,本书的内容还不足以反映我自己在心理学的许多领域内曾经做过的(几乎总是与同事联合做的)大量实验。我最初打算把关于数据的集合论表征的很长的最后一章写成对所指导的实验的复杂活动进行必要的但仍然是希望的一种抽象。最近十年左右,随着计算机能力的迅速增加,许多科学领域的数据分析技术得到了极大提高。我希望在这个方向延伸当前的工作,或许新增加互联网一章是有益的,在今后的十年左右,互联网必定会成为大多数精细的科学出版物的媒介。最后一章的倒数第二节,即 8.6 节,举例说明了我的所思所想。

我感兴趣的另一条进路是,关注许多不同科学思想的历史背景和发展。在哲学中,关注概念和理论的历史发展,有着丰富而迷人的传统。我几乎总是发现,对概率(probability)、物理不变性(physical invariance)、视觉空间(visual space)、心理表征(mental representation)或几乎重要的任何其他科学概念的一种新思想的背景作出分析,即使通常是相当概述性的分析,也是很有启发的和有帮助的。我希望,有些读者对我的许多历史之旅会有同感,这种历史之旅没有对单个概念的演变作出充分详尽的说明。

像这种历时多年才写成的一本书,必然要受惠于更多人对无数论题的修正、远见和建议,这些人多于我可能要明确感谢的人。到目前为止,多数人已经离世,还有许多人已经忘记了他们如何做出了这里所说的贡献。我感谢他们每个人。我一定要提到这样一些人,他们还为我提出了相关评论,或者他们曾一直与我合作完成了在本书的一个或多个观点中曾致谢和用过的

文章。

第二章一开始有斯科特(Dana Scott),后来,还有丘瓦基(Rolando Chuaqui),考斯塔(Newton da Costa),多里亚(Francisco Doria),欣蒂卡(Jaakko Hintikka),塔斯凯(Alfred Tarske),魏因加特纳(Paul Weingartner)和维耶曼(Jules Vuillemin)。第三章涉及的人员有:斯科特,阿罗(Kenneth Arrow),弗勒斯达尔(Dagfinn Follesdal),卢斯(Duncan Luce)和莫斯特林(Jesús Mosterín)。第四章涉及的人员有:卡特赖特(Nancy Cartwright),基娅拉(Maria Luisa Dalla Chiara),德勒斯勒(Jan Drösler)和奥恩斯坦(Donald Ornstein)。第五章涉及人员按字母顺序和年限排列为:布莱克韦尔(David Blackwell),科弗(Thonas Cover),迪亚科尼斯(Persi Diaconnis),法尔马涅(Jean Claude Falmagne),芬斯塔德(Jens Erik Fenstad),法恩(Terence Fine),加拉沃蒂(Maria Carla Galavotti),哈金(Ian Hacking),哈蒙德(Peter Hammond),霍兰(Paul Holland),凯勒(Joseph Keller),卢斯,米勒(David Miller),佩罗-吉马里斯(Marcos Perreau-Guimaraes),波普尔(Karl Popper),罗森克兰茨(Roger Rosenkrantz),鲁阿内(Henri Rouanet)和扎诺蒂(Mario Zanotti)。第六章涉及的人员有:德勒斯勒,因东(Tarow Indow),曼迪(Brent Mundy),奥斯(Gary Oas),潘布奇安(Victor Pambuccian),罗伯茨(Fred Roberts)和鲁宾(Herman Rubin)。第七章涉及的人员有:巴罗斯(Acacio de Barros),法恩(Arthur Fine),加里巴尔迪(Ubaldo Garibaldi),奥斯,圣安娜(Adonai S. Sant'Anna)和扎诺蒂。第八章涉及的人员有:安德森(Theodore W. Anderson),波特纳

xv

(Michael Böttner),埃斯蒂斯(William Estes),韩兵(Bing Han),林良(Liang Lin),吕忠林(Zhong-Lin Lu),佩罗-吉马里斯和威(Timothy Uy)。

我也一直受益于许多届学生的有洞察力的问题与怀疑的评论,这许多年来,他们在课堂上和讨论中阅读了不同的章节。我特别要提到我以前的一组研究生,他们的评论和对错误的纠正是无数的,比如至少有:克兰格尔(Colleen Crangle),德默特尔(Zoltan Domotor),法戈-拉尔若(Anne Fagot-Largeault),霍兰,汉弗莱斯(Paul Humphreys),莱纳(Christoph Lehner),莱文(Michael Levine),曼迪,诺曼(Frank Norman),罗伯茨(Fred Roberts),罗森(Deborah Rosen),罗森克兰茨,斯尼德(Joseph Sneed),齐齐尔夫(Robert Titiev),图奥美拉(Raimo Tuomela)和韦克斯勒(Kenneth Wexler)。

在最后的出版过程中,我必须感谢埃斯科托(Ben Escoto),他最认真地通读了整个底稿,找出打印错误和搞错的地方。也感谢冈德森(Ann Gunderson),他曾勇敢谨慎地制作了有待印刷的正稿。阿里吉(Claudia Arrighi)与我共同整理和校对了大量的参考文献,并作了主题索引。亚当斯(Ernest Adams)阅读了倒数第二章的全部草稿,并且为完善内容和改进文体提出了许多我可采纳的重要建议。像多年来我所拥有的那样,我再一次从他的深思熟虑的批评中获益匪浅。

我着手写作本书时,我还是一位年轻人。对,作为年轻人,当然我至少还不到四十岁。本书完成于我退休十年后的八十岁。当我回过头来再读此书时,我能看到许多地方仍然有待完善,也许,最重要的是补充细节和更仔细地检查差错。但我知

道，现在该是到了收笔的时候了，而且我也这么做了。

我把本书献给我的五个孩子。在他们的整个生活中，我一直在撰写本书，老大帕特里夏(Patricia)除外，不过，在我开始撰写时，她还很小。我也要表达对我的妻子克里斯蒂娜(Christine)的感激之情，她耐心地容忍了我多年来断断续续的努力，现在终于完稿了。

帕特里克·苏佩斯

斯坦福，加利福尼亚

2002年3月

目　录

1.
绪　论

1.1　一般观点

1

根本没有一种简单的或直接的方式来刻画科学哲学的特征。即使是很狭义的和更明确的形式逻辑学科,也不是一门真正精确定义的学科。个别哲学家的科学哲学观变化很大,而且在关于什么是最重要的主题上分歧很大。尽管如此,关于像因果性、归纳、概率和理论结构之类的特定主题有着相当广泛的共识。在本书中,我用特定的形式方法接近这些共识和相关主题试图表明,我们如何能够用这些方法使在一般论文中很容易被忽视的区别变得更加明确。另一方面,我的目标是不让形式问题变成为晦涩难懂。我一直设法沿着一条小路,使读者不会在纯技术问题的丛林中迷失方向。

有人可能会认为,强调形式方法来自这种愿望:着重讨论在通常能给出数学表达的科学中特别成熟的那些理论。毫无疑问,这种考虑在某种程度上是重要的,但不是最重要的。从我自己的观点来看,强调形式方法在科学哲学的系统讨论中的作用的一个更基本的特征,有两种理由。一种理由是,希望对可以用

来整理和批评现有学说的一个固定的参考框架或一种确定的一般方法作出任何大规模的讨论。形式的、集合论的方法为讨论系统的科学哲学问题提供了这样一个一般的框架。这些方法在这里比在数学基础中更适当,因为集合论的基础本身是数学基础研究的一个中心主题。把数学基础问题与科学基础问题区分开来,似乎是一种明智的劳动分工。据我看来,科学哲学的大多数核心问题,能够在不质问数学基础的前提下,通过接受像集合论的标准表达那样的方法,作出很详细的讨论。我在讨论科学哲学问题时,把形式方法等同于集合论方法。我在后面更详细地追踪这种等同性的理由,但我确实希望明确的观点是,我没有为这种等同性的最终特征,乃至集合论进路的最终特征,提供任何教条的承诺。从我在第八章关于语言的行为与神经理论的论述中可以看出,我自己关于这些现象的最终令人满意的理论观,将会超越标准的集合论进路。

2

无论如何,集合论进路的一个优势是,我们可以很容易地应对科学哲学中通常引入的人工语言经常受到的批评——即这样的语言不足以有力地表达大多数科学结果。我们所运用的集合论的方法与框架,很容易有力地表达任何一门经验科学中或科学的一般逻辑中的任何一种系统化的结果。

在科学哲学中提倡形式方法的另一个理由是,深信对证据问题的常识性处理和人工语言的处理都是不充分的。这两条进路为在经验科学中评价证据的异常复杂的和技术上的实践问题,提供了实在过分简化的说明。较长的第五章关于概率的一些部分举例说明了许多相关的敏感问题。

在转向详细地考虑集合论方法的第二章之前,在这一章绪论中对科学理论进行非正式的讨论将是有用的。这种讨论旨在勾画出后面要更彻底地考察的许多问题。

1.2　什么是科学理论?

当我们通常询问什么是某某时,我们期待着能得到一个清楚而明确的回答。例如,如果有人问我,什么是有理数,我可以给出一个简单而精确的回答:有理数是两个整数之比。还有一些其他类型的简单问题,也能给出精确回答,但是对于这些问题而言,通常给出和接受相当模糊的回答。某人在一本书中读到了油桃,但他从来没有见过油桃,或者可能看见过油桃,但不熟悉油桃的英语名称。他可能问我,"什么是油桃?"我可能答复说,"一种皮很光滑的桃子"。当然,这不是一个很确切的回答,但如果我的提问者知道什么是桃子,这种回答可能几乎是令人满意的。"什么是科学理论"的问题与这两种模式都不相符。科学理论与有理数或油桃不一样。无疑,科学理论与油桃不一样,因为它们不是物体。它们与有理数一样,都不是物体,但它们又与有理数完全不一样,在这方面,科学理论不可能根据其他非物理的抽象对象作出简单的或直接的定义。

"什么是物理学?""什么是心理学?""什么是科学?"这些熟悉的质问提供了相关问题的很好事例。我们确实无法对这些问题给出简单而精确的回答。另一方面,对物理学或心理学这些问题能够做出许多有趣的评论。我希望表明,科学理论也是如此。

传统概述。科学理论的传统概述——我强调了"概述"这个词——有点像下列方式进行的:一个科学理论是由两部分组成的。一部分是抽象的逻辑演算,这种演算包括逻辑学的词汇和该理论的原始符号(primitive symbols)。用理论的原始符号陈述理论的公理(axioms)或公设(postulates),来确定理论的逻辑结构。对于许多理论而言,原始符号被认为是像"电子"或者"粒

3

子”之类的理论术语,这些术语不可能以任何一种简单的方式与可观察的现象相关。

理论的第二部分是这样的一组规则,这些规则通过为最起码的某些原始的和确定的演算符号提供通常所谓的“操作定义”(coordinating definitions)[①]或“经验解释”(empirical interpretations),把经验内容赋予逻辑演算。人们总是强调,只有第一部分不足以给科学理论下定义;因为如果不对理论预期的经验解释作出系统的详细说明,就无论如何不可能将理论作为科学的一部分对之进行评价,尽管它能作为纯数学的一部分得到简单的研究。

这种描述最突出的东西是它的高度的示意性(schematic nature)。就理论的第一部分而言,在科学哲学家的著作中,把一个理论实际上当作逻辑演算的事例几乎是罕见的。许多模糊的论证过分沉溺于证明:作出逻辑演算原则上是简单的,而且只是一个单调乏味的细节问题,但是很少给出具体的证据。理论的第二部分的概述,也就是说,某些术语的操作定义或经验解释也是高度示意性的。对所提供的相对模糊的框架作出的一种常见辩护是,各种不同的经验解释,例如,测量质量的许多不同方法,使精确描述很困难。此外,当我们从精确地阐述的理论转移到几乎所有的科学家都使用的很不确切和简洁的实验语言时,很难把一种定义模式强加于经验解释的规则。

我想要支持的观点,不是说这种标准的概述是断然错误的,而是说它过分简单。它的概述性使它有可能既遗漏了理论的重

① “coordinating definitions”从字面上应翻译为“协调定义”,但是我在与作者讨论这个术语的准确译法时,他提供了一个替换词,“operational definitions”,于是,这里根据这个替换词,译为“操作定义”更易于理解。——译者

要特性，也遗漏了理论之间可能引入的重要区别。

理论的模型与经验解释。 首先，许多哲学家表现出一种很强的倾向，即把理论的第一部分说成是完全用句法术语进行的纯逻辑演算。第二部分提供的操作定义，在现代逻辑的意义上，没有为形式演算提供一种适当的语义学。且不说关于直接的经验观察的问题，从逻辑的观点谈论理论的模型是恰当的和自然的。这些模型是抽象的、非语言的实体，它们的概念通常远离经验观察。因此，除了逻辑之外，有人可能有理由问这样的问题：一个模型概念能对常见的理论的经验解释的讨论增加什么呢？

我认为，大多数哲学家发现，谈论理论比谈论理论模型更容易，这种说法是正确的。其理由有几种，但也许最重要的两种理由是：第一，哲学家的理论范例通常在特征上是相当简单的，因而易于以直接的语言方式进行讨论。第二，引入理论模型必然会把一个较强的数学元素引入到讨论当中。当理论以所谓标准的形式化给出时，作为语言实体来谈论理论，即明确地说出理论的精确定义的句子集合等，是一件很自然的事情。当我们在一阶逻辑的范围内阐述理论时，通常说理论有一种标准的形式化。大体上说，一阶逻辑正好是为一类对象提供了句子连词和谓词的逻辑。不幸的是，当一个理论假定了超出一阶逻辑的范围时，还以这种方式阐述理论就既不自然，也不简单。例如，如果在公理化的几何学中，我们希望把线定义为是某些点的集合，我们必须在已经包含有集合论思想的框架内进行。诚然，对几何学和集合论的相关部分同时公理化在理论上是可能的，但这是非常棘手的和很困难的。更复杂的结构的理论，像量子力学、经典热力学或现代定量的学习理论，不仅需要运用一般的集合论思想，而且需要运用关于实数的许多结果。在一阶逻辑的范围内对这些理论进行形式化是完全不切实际的。这类理论在其复杂性上

4

非常类似于纯数学中主要研究的理论。在这样的语境中，更加简单的是断言关于理论模型的问题，而不是直接地和明确地谈论理论的语句。也许这方面的主要理由是，当理论不以标准形式化出现时，理论的一个语句的概念是很不明确的。

我喜欢举两个例子，在这两个例子中，模型概念以自然的和明确的方式研讨科学理论。第一个例子是关于测量本性的。特定的测量理论的基本目标是以精确的方式表明，从定性的观察如何过渡到更精致的科学理论阶段所需要的定量的断言。对实验上可实现的运算和关系的适当代数的公理化提供了如何实现从定性到定量这个过渡的一种分析。已知某一经验参量的一个公理化的测量理论，例如，质量、距离或力，数学的任务是为该理论的**模型**证明一个表征定理，粗略地说，它确定了任何一个经验模型都同构于(isomorphic)该理论的某种数值模型(numerical model)。[①] 模型之间存在的这种同构性(isomorphism)证明了把数(numbers)应用于事物(things)是正当的。我们确实无法把手头的一个数应用于一个物体。我们所能做的是表明，经过某些经验操作所产生的一个现象集的结构，与经过算术运算和关系得到的某种数集的结构是一样的。模型在给定语境中的同构性定义使同样结构的直觉想法变得更加精确。发现模型的这样一种同构性的重大意义在于我们就可以把所有我们熟悉的计算方法的知识应用于算术模型，推论出关于同构的经验模型的事实。对与数值模型同构的一个测量理论的经验模型这个核心概念给出一种语言表述，是极其困难的和乏味的。但是，用模型

5

---

① 第三章阐述表征定理的概念。在这种语境中，同样重要的是承认，这里所用的一个经验模型的概念本身是从实际测量的经验过程的大多数经验细节中抽象出来的。经验模型的作用是以一种系统化的方式整理所运用的测量程序的**结果**。我在下面关于操作定义和理论的层次结构部分进一步评论这个观点。

论的术语来说,这个概念是简单的,事实上,正如第三章所表明的,这个概念是纯数学的所有领域用到的很一般的同构表征概念的一种直接应用。

运用模型的第二个例子涉及科学哲学中对还原论(reductionism)的讨论。与把一门学科还原为另一门学科的问题相关所阐述的许多问题,可以作为运用一个理论模型的表征定理的一系列概念问题来阐述。例如,对于许多人来说,针对一个心理学理论的任一模型,都有可能在某个生理学理论中建构一个同构模型,表明这一点会适当地确立可以把心理学还原为生理学的论点。当前,不论是在心理学领域,还是在生理学领域,都缺乏任何深层次的统一理论,这使得现在试图解决这样一个还原论问题毫无希望。来自物理学的经典例子是把热力学还原为统计力学。尽管从逻辑的观点来看,这种还原通常并没有以绝对令人满意的形式得到过阐述,但毫无疑问,它在本质上是正确的,并代表了经典物理学的一个伟大胜利。

第八章的两个密切相关的学习理论证明了一个真正的还原定理。即使这种在概念上相当简单的情形,也要求进行大量的技术论证。在数学上提供从热力学到统计力学的可接受的还原的困难,是难以克服的,而且基本文献中已经认可了这一点(例如,参见,Khinchin, 1949; Ruelle, 1969)。

理论的内在描述与外在描述。除了刚才提到的理论模型概念的两种应用,我们可以把理论的模型概念直接应用于描述科学理论的问题。我希望在内在特征和外在特征之间作出对比。把一个理论作为一种逻辑演算的表述,或者用我更喜欢的术语来说,作为具有标准形式化的一个理论,提供了一种内在特征,但这当然不是惟一的进路。例如,在逻辑语境中很自然要问的问题是,是否能够根据标准的形式化,即在一阶逻辑的范围内,

使一个特定的理论公理化。为了精确地阐述这样一个问题，有必要用某种外在方式来描述理论。提供这样一个外在描述的最简单的方式之一是，只定义该理论的预期的模型类。问我们是否能把这个理论公理化，恰好是问我们是否能陈述一组公理，使得这些公理的模型恰好是所定义的类中的模型。

作为在外在意义上和内在意义上表述理论的一个很简单的例子，考虑在熟悉的小于关系条件下，与一个实数集同构的简单排序(orderings)理论的外在表述。也就是说，考虑与实数的小于关系的某个片段同构的所有二元关系的类。一个理论的外在特征通常遵循为这些排序所规定的类型，即我们指定一个特殊的理论模型(在这种情况下是数值上的小于关系)，然后描述与这种独特模型相关的整个理论模型的类。内在特征的问题现在是表述一个公理集，该公理集在不涉及模型之间的关系只涉及任何一个模型的内在特性的前提下描述该模型的类。根据当前的情形，这种解决方案是相对简单的，尽管这还不是在一阶逻辑的范围内进行的自然表述。①

对科学理论的偶尔审查间接地表明，通常的表述在特征上是内在的，而不是外在的，因此，外在表述的问题通常只出现在纯数学中。这似乎是一种令人愉快的结果，因为我们的哲学直觉的确是，一个内在特征一般说来比一个外在特征更可取。

然而，一个科学理论的内在公理化的问题会比这一评论所表明的更加复杂和更加敏感得多。幸运的是，恰好通过对理论模型类的明确考虑，就能正确地看待问题，并以有可能考虑其精

① 内在公理恰好是，一个简单排序加这种排序必须在其定义域内包括一个可数子集的公理，稠密(dense)就与所讨论的排序相关。第三章从形式上给出这个例子的细节。

确解的方式表述问题。在这一点上,我仅概述性地举一个简单的例子。经典质点力学的公理通常是根据默许地假定的作为一个参照系的坐标系来陈述的。这种结果之一是,从公理中推论出的关系,相对于伽利略变换,未必是不变的。我们能够把一个参照系的默认假设看成是理论常见特征的一个外在方面。从理论模型的观点来看,对力学进行标准公理化的困难在于,可以用大量形式上不同的模型来表达同样的力学事实。每一个不同的模型都代表了对一个不同参照系的默认选择,但是表征相同力学事实的所有模型都是通过伽利略变换相关的。因此,公平地说,在这个例子中,通过伽利略变换相关的模型之间的不同没有任何理论意义,可以把这看成是下列公理的一个缺陷:即,这些有着细微差别的不同模型是存在的。重要的是意识到,关于通过伽利略变换相关的模型的这种观点,不是在理论的经验解释标题下通常提出的那种观点。这个概念要点恰好属于物理学的理论问题。我在这里介绍这个例子是为了提供这样一个简单的事例:对模型的明确考虑如何能使关于科学理论本性的讨论更加微妙。从哲学的观点来看,一定有可能坚持认为,质点力学,作为一个科学理论,应该只用伽利略的不变关系来表示,而且惯用的表述在这方面是有缺陷的。第六章较详细地讨论了这些问题;第四章阐述更一般的不变性理论。

操作定义和理论的层次结构。我现在转向上面提到的理论的第二部分。在前面的讨论中,我们用"理论"这个词只指理论的第一部分,也就是说,指理论的公理化,或者作为一种逻辑演算的理论表达;但正如我在一开始所强调的那样,提供一个理论的经验解释的必要性与阐述这个理论的形式问题恰好是同样重要的。我关于理论的这方面的核心观点是,这个故事比关于理论的操作定义和经验解释的熟悉评论所表明的要复杂得多。哲 7

学家通常描述的这类操作定义，在流行的理论的哲学阐明中占有一席之地，但在检验科学理论的实践中，为了把理论与数据联系起来，要求作出更详细的阐述和提供更精致的形式机制。科学家标明一个实验的具体经验本身，不可能在任何完备的意义上与一个理论相联系。经验必须经过概念加工，这一点在许多情形下是很不精致的，所呈现出的是标准形式的实验数据。这些标准数据建构了实验结果的一个模型，而且直接的操作定义是为这个模型提供的，而不是为理论的模型提供的。还有一个特征是，实验结果的模型与理论的任何一个模型都具有相对不同的逻辑类型。一个理论的模型通常包含有连续函数或无穷系列，但实验结果的模型在特征上是高度离散的和有穷的。

对实验结果的模型和某个指定的理论模型关系的评价，是近代统计方法论的一个典型的基本问题。从当前的目的来看，这种方法论的重要之处在于：首先，它本身在本质上是形式的和理论的；第二，这种方法论的一个典型功能一直是要提供一个详细阐明的实验理论，它能在任何一个基本的科学理论和原始实验经验之间作出调解。我这里的论点只是，使这种层次结构的存在变得更加明显，并且指出根本没有一个简单程序能给出理论的操作定义。认为在刚提到的数据的标准形式的意义上，给出操作定义是确定理论模型与实验结果模型之间的适当联系，这种说法简直是任意删改事实。例如，参考操作定义不足以涵盖根据这些实验结果的模型来评价理论模型中的理论参量的详细方法。①

① 提出实验程序的模型，不只是实验结果的模型，也会是令人向往的。在这个方向上一个很详细的步骤是，一定会用心身概念和相关的心理学概念来描述实验科学家在他们的实验室里实际上在干些什么。这里不对这个重要的基本主题作出阐述，而且在科学哲学文献中也很少有系统的阐述。

如果有人问"一个科学理论是什么?"似乎对我来说,根本给不出一个简单的回答。我们将会把为了检验理论而精心构想的统计方法论作为理论的一部分包括进来吗？如果我们会认真地接受这种标准主张：操作定义是理论的一个组成部分,那么似乎不可避免的是,我们也一定在更详细的理论描述中包括了设计实验、评价参量和检验理论模型适合度的方法论。似乎对我而言,给出下列形式的精确定义是不重要的：X是一个科学理论,当且仅当,如此这般。重要的是承认,存在着由检验基本理论的实验方法论所产生的理论的层次结构,这是任何一门精致的科学的学科的一个基本要素。 8

本章后面的部分没有系统地阐述检验理论的统计方法论的重要主题。第五章关于概率的解释或表征,即本书最长的一章,是统计方法分析的一个详细绪论。我在几本较早的出版物中已经讨论过这些问题,[①]第六章对视觉空间的本性进行了相当详细的实验讨论,第八章的最后一节根据极值统计概念的经验应用对词和句子的脑电图表征的数据进行了相当详细的讨论。

工具性的理论观。我迄今还没有提到无疑是相当重要的一种科学理论观;这种观点是,从工具性的观点来看待理论。按照这种观点,理论的一个最重要的作用不是整理陈述或断言陈述的真伪,而是提供可以用来从一组事实推论出另一组事实的重要的推理原理。这样,在熟悉的三段论的推理中,"所有的人都是会死的;苏格拉底是人;因此,苏格拉底是会死的",按照工具性的观点,大前提"所有的人都是会死的",被转化为一个推理原理。于是,在这个三段论的推理中,现在只有小前提"苏格拉底

---

① Suppes and Atkinson 1960, Ch. 2; Suppes 1962, 1970b, 1973a, 1974b, 1979, 1983, 1988; Suppes and Zanotti 1996.

是人”。从逻辑的观点来看,显然这是相当不重要的步骤,而且,很自然地提出的问题是,就工具性的观点是否还有更重要的东西有待阐明呢?或许,对于在看待理论或定律的这两种方式之间不仅仅存在着言语差异这种主张的最有趣的论证是:当把理论看成是推理原理而不是大前提时,我们不再直接关心确定理论的真伪,而是评价理论在推断新的事实陈述时的有用性。最初的形式概念实际上没有一个概念产生于这些哲学讨论:它们能替代真理和有效性的经典语义概念。比如说,除非引入某些系统的语义概念来代替这种标准分析,否则,谈论定律与事实陈述具有不同的职责,是没有价值的。

从另外一个方向来看,有一种认真的具体努力是,提供一个评价理论的形式框架,它可以取代经典的真理观。我想到了近代统计决策理论。典型的统计决策理论是谈论行动,而不是陈述。一旦焦点从陈述转向行动,似乎就相当自然地用预期的损失或风险概念取代了真理概念。问一个陈述是否正确,是适当的,但是,问一个陈述是否有风险,是没有多大意义的。另一方面,问一种行动有多大风险,而不是问这种行动是否正确,是合理的。显然,照字面意义来理解的话,统计决策理论比已经概述的观点更根本地突出了理论的工具性观点。理论也不能被看成是推理原理,而应被看成是整理证据的方法,以决定在几种行动中采取哪一种行动。当把理论看成是推理原理时,一个直接的问题是,返回到经典的观点,并把作为推理原理的理论与在一个论证中作为一个真的大前提的理论概念联系起来。经典的观点与作为导向采取一种行动的工具的理论概念之间的联系,肯定更加疏远和间接。

9

尽管在有关统计学基础的文献中,应用统计决策理论思想的许多事例是很成功的,但是这些事例决不与复杂的科学理论

相关。再一次公平地说，当我们想谈论关于精致的科学理论的评价问题时，像统计决策理论那样的学科迄今还没有提供任何一种真正的选择来替代真理和有效性的语义概念。事实上，即使是对统计决策理论的偶尔审查也会表明，尽管基本思想有工具性趋向，但是，理论的形式发展完全取决于标准的语义概念，而决非取而代之。我这里的意思是说，当把关注集中在作为探究的最终状态的作出行动时，决策理论家已经发现，在描述证据、他们自己的理论等的过程中，有必要运用标准的语义概念。比如说，我记不得决策理论家进行过这样一次讨论，在这种讨论中，他们是根据效用，而不是根据其真伪，来处理特殊的观察陈述。

目前，统计决策理论似乎显然没有提供一种真正一致的或完全原创的新的科学理论观。也许，决策理论的未来发展将会向着这个方向进行。不管怎样，还是有我喜欢作为本绪论的最终要点来讨论的一种更根本的工具性观点。正如我已注意到的那样，许多工具性分析的特征是，把理论的状况与对事实的特殊断言的状况区分开来。运用语言的更激进的工具性的、行为主义的观点，向这种区分提出了挑战，并且从一种行为主义的观点考虑整个语言的运用，包括理论陈述和特殊的事实问题在内。按照对这个问题的这种观点，对语言的所有用法进行分析都将尤为强调语言的使用者。据说，现代逻辑的语义分析远不足以对语言的认知用法给出说明，因为它没有明确地考虑说话者、作者、听众和读者对语言刺激的产生与接受。很明显，对于行为主义者而言，根据前面考虑的那些概念，他们无法对“什么是科学理论”这个问题提供最终有意义的回答。一个适当而完备的回答，只能根据对理论的生产者和使用者的明确而详细的考虑来提供。一般情况下，用这种行为主义的方式考虑理论或语言是

10 很有吸引力的。然而,当前它所缺乏的是充分的科学的深刻性和明确性,据此它可以作为一种真正的选择来代替现代逻辑和数学的进路。此外,前面所讨论的理论与模型的大多数语言确实是如此近似正确,以至于对我们看待理论的方式作出任何行为主义的修正,都一定会把模型和理论作为一级近似来谈论。未来的问题是,看看行为主义的进路是否会加深我们对科学理论本性的理解。第八章的结尾,即本书的最后一节,考虑了某些新的方向,因为从行为主义的心理学转向了作为把语言分析扩展到其脑电图表征的适当框架的认知神经科学。

这种工具性的科学观的一个不同方面是它与实用主义有密切的联系,尤其是,为了达到某种别的目的,最好是实践的目的,只认真地接受有用的东西,在这种更根本的意义上,与实用主义的联系更密切。对于概率论的这种实用主义的态度,第五章的最后一节作出了某种详尽的考虑,正如所评论的那样,这一节完全专注于概率的表征问题。

1.3 本书的规划

古希腊人最非凡的、也许是最杰出的智力成就之一,是他们明确地发展了公理的分析方法。欧几里得(Euclid)在公元前300年撰写的《几何原本》可能是科学史上最有影响的一部著作。每一位受过教育的人都知道欧几里得的名字和他大致所做的工作——他以一种系统化的方式以一组公理和几何假设阐明了几何学。

第二章的目的是详尽地阐述近代的公理化方法的概念。除了有许多不同的形式细节之外,背离欧几里得是不过分的。这一章关于近代公理化方法的中心论点是表明,在集合论的范围

内,如何有可能把任何一个数学分支或科学理论公理化。尽管用了简单的数学事例来举例说明关键的概念要点,但是这里关注的是科学理论,不是数学部分。在集合论的范围内把一个科学理论公理化,是使其结构变得精确和清楚的重要的第一步。一旦提供了这样一种公理化,就有可能问近代数学特有的这类结构问题。例如,一个理论的两个模型什么时候是同构的呢?也就是说,这两个模型什么时候会精确地拥有相同的结构呢?首先,第二章明确表明,在集合论的框架内,如何使一个理论公理化,就是像我多年前提出的那样,**定义一个集合论的谓词**。

在第二章探讨了公理化方法之后,接下来的两章是讨论表征问题(第三章的中心论题)和不变性问题(第四章的焦点所在)。这里,我只想谈这几章的要点。

最重要的区别是,我们在谈论一个已知理论的表征时,不是谈论该理论本身,而是谈论该理论的模型。当一个理论的任何两个模型是同构的这一特殊情形实现时,那么这个理论被说成是范畴的(categorical)。在标准的阐述中,实数理论就是这样一个例子。另一方面,几乎没有任何一个科学理论在特征上应该是范畴的。当一个理论不是范畴的时,一个重要问题是发现, 11 能否找到该理论的一个令人感兴趣的模型子集,使得任何一个模型与这个子集的某一项同构。寻找一个理论的这样一个非凡的模型子集,并且表明这个子集具有所指明的特性,就是证明该理论的一个**表征定理**。第三章的目的相当广泛地讨论了表征概念,而且最初并不只限于理论模型。一般的表征概念目前在哲学中很流行,因而人们会发现,在文献中对表征含义有许多不同的描述。我不打算如何详细地涵盖这种广泛的可能性,而是集中于这样的问题:我为什么认为刚才以一种非形式的方式描述的集合论的同构表征概念是最重要的概念。

第四章致力于更微妙的不变性概念，列举了不变性的许多数学例子，不过，在科学理论中，突出不变性定理的两个标志性领域是测量理论和物理理论。这里是一个简单的测量例子。例如，如果我们有一个测量质量的基本理论，那么满足这个理论定理的任何一种经验结构，都与在这个经验结构中表示物体的数值测量的数值模型是同构的，或者有点接近于同构的。现在，正如我们大家知道的，这些测量通常是用质量或重量的特殊单位（比如，克、公斤或磅）来表示。分析不变性的观点是表明，从结构的观点来看，通过改变测量单位（测量单位是任意的，没有反映任何本质的东西），与第一个数值模型相关的任何其他数值模型，仍然是一个令人满意的数值模型。这似乎很简单，但是，当详述各种不同的测量理论时，问题就变得相当错综复杂。

思考不变性的另一种方式是根据对称性。正如我在第四章开头所表明的，对称性和不变性是密切相关的。一个正方形围绕其中心旋转 90 度，其形状保持不变，是因为正方形有明显的对称性。一个三角形，不是正方形，没有这种对称性，只在旋转 180 度的条件下才是不变的。

我提到的第二类不变性是特殊的物理理论的不变性。也许，最熟悉和最重要的例子是与狭义相对论相联系的不变性概念。就物理学家易于阐述的问题而言，研究其运动速度接近于光速的物理现象的现代物理理论，必须在相对论的意义上是不变的。我在这里不想精确地说明在这一点上是什么意思，不过，除了第四章的一般阐述之外，第六章还对这些问题进行了相当彻底的讨论。

值得关注的是，表征的许多哲学讨论，似乎并不关注不变性问题。但是，在能够证明表征定理的那类明确的科学理论中，不变性概念被认为是至关重要的。第四章的目的之一是尽力说明

情况为什么会如此。

在第五章，我也许会转向一般科学方法论中的核心的哲学主题：即概率的本性。概率没有进入到许多科学理论当中。最 12 重要的事例是经典物理学的理论，其范围是从牛顿的质点力学理论到流体力学理论、柯西的热理论，再到麦克斯韦的电磁现象理论。但是，当人们承认在检验这样的理论时必须以某种方式说明测量误差时，概率在 18 世纪和 19 世纪就已经是重要的。关于观察误差的概率分析的奠基性文集是由辛普森(Simpson)、拉格朗日(Lagrange)、拉普拉斯(Laplace)、高斯(Gauss)等人撰写的。从那时起，特别是大约从 20 世纪初开始，概率和统计已经成为所有关于科学方法论的详细讨论的主要部分。

尽管人们在科学方法论中承认概率的重要性，但是关于概率的本性还远没有达成共识。第五章的目的在于为一直被坚持的最突出的观点提供一种从容不迫的分析。为了尽可能地给出这种讨论的焦点，我阐述了从拉普拉斯的概率概念到柯尔莫哥洛夫(Kolmogorov)的复杂度观点(某种表征定理)的每一种特有见解。哲学讨论通常伴随着在形式上呈现出许多资料，从而使这一章成为本书最长的一章。

概率在科学方法论和科学理论阐述中的重要性已经得到了广泛的认可。现在，它是许多科学领域内全面阐述理论的一个基本概念。20 世纪最重要的科学领域是统计力学和量子力学，第七章讨论它们的某些基本特征。

因为某些感兴趣的读者不熟悉从柯尔莫哥洛夫在 20 世纪 30 年代提出的经典公理中发展出来的标准形式的概率理论，所以，我在转向对概率本性进行的任何一种不同的表征或解释之前，在第五章的开头对此给出了一个详细而基本的阐明。

最后三章研究特殊学科中的表征和不变性问题。第六章和

第七章聚焦于物理学和心理学。最后一章即第八章,集中于语言学、心理学,最后是神经科学。

第六章专门关注空间与时间的表征。第一节研究经典时空,第二节研究相对时空(在狭义相对论的意义上),主要结果关涉不变性;这种表征或多或少是假定的。这是因为在经典物理学和相对论物理学中,不变性问题是至关重要的。6.4—6.7节转向视觉空间中的不变性问题,更一般地说,转向知觉的不变性问题,因此,在那里关注的是心理问题而不是物理问题。在这几节提出的问题更加特殊,它们很少被认为是永恒的问题,但却反映了我自己的相当特殊的兴趣。进行这样一种推广处理的另一个理由当然是这样的情形:科学哲学家通常对经典物理学和相对论物理学中讨论的主题,比在分析视觉空间时所用的实验与概念更熟悉。最后一节是关于几何学中的有穷论(finitism)问题,即一个在古希腊的几何学中有其根源,还与更一般的数学基础与应用中的有穷论相关的主题。

13

第七章转向力学中的表征问题。我原来打算这一章完全集中于我最早的兴趣之一,经典质点力学的基础。这里为了在某种程度上对这些公理给出稍微不同的阐述,我喜欢重写这些古老的基础。但是,当我反思物理学中的表征的本性时,我忍不住补充了某些部分来涵盖更新近的论题。我这里首先想到了许多人在量子力学的隐变量问题上做的大量工作。几年前,我与两名物理学家阿卡西奥·德·巴罗斯(Acacio de Barros)和加里·奥斯(Gary Oas)总结出许多关于隐变量的定理和反例。对那件工作作出描述似乎是适当的,因为在阐述这些结果时包括了许多不同的表征定理。我也把巴罗斯和我自己最近的工作包括进来。远距离纠缠和非定域性的微妙问题与当前不存在定域隐变量的问题相关,在我看来,这是当前量子力学提出的最令人

困惑的哲学难题。我还在最后一部分包括了因果性过程中的可逆性问题。书面的进路既含有决定论的过程,也含有随机过程。第五章介绍的概率概念扩展到涵盖时间的过程,特别是动力学的过程。我认为,我提出的弱可逆性与强可逆性之间的区别,似乎没有在关于可逆性的哲学讨论中得到应有的强调,尽管这种区别有时在关于随机过程的文献中被用到。无论如何,对这些问题的思考再次导向了关于表征,特别是关于所有类型的因果性过程的不变性的有用结果。

最后一章,即第八章,是对多年来所阐明的语言表征的各种结果的一种大杂烩。这反映了我自己在心理学和语言学领域内的交叉兴趣。这一章首先评论了诺姆·乔姆斯基(Noam Chomsky)、斯蒂芬·克伦(Stephen Kleene)等人关于指定类型的文法和指定长度的自动机(automata)或计算机之间具有的共同的表征定理的漂亮的经典结果。这些结果,现在已有四十多年,无疑还会在很长时间内是关于这些问题的部分永久文献。概念与定理的整理来自我在20世纪60年代和70年代多年来讲授自动控制理论课的讲稿。在关于这些结果的这几节之后,我转向了我自己根据刺激-反应模型论述自动机表征的更特殊的工作。我认为,搞明白如何从起源于经典行为主义但需要用数学表示的简单的心理学思想转变为自动机的表征,并由此得出相对应的文法,是有用的。接下来的一节完全聚集于为只有反映和强化概念的行为学习模型证明一个表征定理。只给定这两个概念,一种反映的概率依赖于整个过去的强化史。为了切掉对过去的这种很强的依赖性,用扩展到刺激及其条件作用(或联想)的理论模型来提供实际上是马尔可夫(Markov)提出的表征手段,即把对过去的依赖限于前面的试验。下一节(8.5)延续到我近年来关于自然语言的机器学习所做的某些详细的工作。

14 这里，精心提出的基本公理表达了关于机器（计算机或机器人）如何能够通过联想、抽象和概括来学习自然语言片段的合理的心理学思想，但却是相当明确的思想。最后，在结束这一节时，我把我最近关于词和句子的脑电图表征的某些工作列了进来。在提出一个作为大脑加工处理的词的试探性表征时，把经验工作和概念工作结合起来是一个结束的好地方。我在前面几页已经分析的许多结果没有永恒性，但这是在当前科学的框架内提出有用的表征和不变性思想的一个无法拒绝的机会。

1.4 如何阅读本书

正如前面我在描述本书的各章时表明的那样，细节和技术难度是变化很大的。因此，我在这一部分试图提供某种导读。我是为几种不同类型的读者写的，不仅包括科学哲学家，而且也包括对基本问题感兴趣的不同学科的科学家。许多概念和理论的历史背景是概述性的。这符合我自己了解一些给定理论的前期发展的爱好，即使其详细程度离提供一个详实的历史还相差甚远。

我的意图是，从第一章到第四章对本书的主要论题，特别是表征和不变性论题，提供一个合理的初步介绍。这种主张有两个例外。第一，在 3.5 节中，部分递归函数（即可计算的函数）的机器表征在风格上是相当技术性的，即使原则上对每一个细节都作出说明。本书后面的不同论点运用了一般的表征结果，但随后不需要对此作出证明。第二，第四章的最后一节，关于熵在各态历经理论中的作用，要求有某些特殊的背景知识才能完全理解，但这个不变性的例子是如此的漂亮与美妙，使我忍不住把它包含进来。我还要强调指出，这些中心定理都没有加以证明；

本书只对它们进行了谨慎的阐述与说明。读者可以在不失连贯性的前提下跳过 4.5 节。第一章到第四章的另外一点是，有一些证明只限于脚注，3.5 节除外，或者在 3.4 节的情况下放到这一节的最后。

正如我所说的那样，我也许把概率算作是科学哲学的一个最重要的概念。但是，我在这里面临着向一般读者作出阐明的问题。除了第四章的最后一节之外，第一章到第四章没有详细地讨论概率。这样，对介绍概率的基础感兴趣的一般读者应该如何阅读第五章呢？首先，不熟悉现代概率的形式概念(比如，随机变量的核心概念)的人，应该阅读 5.1 节中对形式理论的非常详尽的介绍。已经熟悉形式理论的那些人，应该跳过这一节。真正的问题是如何处理其余的几节，包括含有许多技术细节的几节在内，特别是关于概率的倾向性表征的很长的一节(5.6)和关于主观概率观的同样长的下一节(5.7)。我建议，希望大致了解情况的读者，可阅读第一节之后的其他每一节的非技术性部分。证明等很容易被省略掉。对于某些读者有用的另一个策略是，浏览论及不太熟悉的概率观的部分来决定这些部分是否值得更详细的一读。 15

最后三章涵盖了许多论题。幸运的是，这几章的大多数内容几乎是彼此独立的。例如，对知觉和心理学概念感兴趣的人可以阅读 6.4—6.7 节，用不着阅读关于经典时空和相对论时空的 6.2 和 6.3 节，主要对物理学基础感兴趣的人，则反之。

第七章的每一节都与物理学相关，尽管关于可逆性的最后一节(7.3)有更广泛的影响，可应用于生物学和社会科学研究的许多过程。另一方面，在第八章中，实际上没有一节与物理学基础方面的中心论题相关。焦点始终是关于语言的，但同样或也许更多的是心理学问题，而不是语言学问题。

我这里再次谈一下我在绪论中所说的本书最后的定理汇总表。这个表提供了各章明确阐述(但不必证明)的表征和不变性定理的一个列表。根据这个表能很快找到一个特殊定理的出处。

像战火和恋情一样,在哲学问题与科学问题中,简明扼要地浏览是重要的。我们大家都没有时间去察看让我们感兴趣的所有细节。当我们从表面上精读不熟悉的资料时,我们从好奇之处获得了新的思想,而且,我们通过联想很少会意识到这种新思想。如果本书的某些读者会是这样,我将深感欣慰。

2. 理论的公理化定义

本章首先详尽地考察了 *Model* 这个词的某些含义，这些含义可以从其用法中推论出来。① 第二节在一阶逻辑范围内简明扼要地阐述了理论的形式化，考虑了这些理论模型的事例。第三节是本章的核心部分，阐述了用集合论谓词定义的科学理论的公理化特征。于是，关于科学基础的这条进路就与第四节的公理化方法的更悠久历史联系了起来。后面几章提供了把理论公理化的大量事例。 17

2.1 *Model* 在科学中的含义

"*model*"这个词的使用不只限于科学语境中。它被用于所有类型的普通场合。每个人都熟悉建筑物、船和飞机的物理模型概念。确实，这样的模型经常被作为礼物送给小孩子，特别是飞机模型和车模。(这种用法在工程中还有广泛的技术应用。)在日常谈话中，例如，在已经满足某种设计标准的地方，也提到模范兵团、模范政府、典型条例等。在许多情况下，例如，谈论模范政府，并没有指向任何一个实际政府，而是指感觉一个理想政府应该满足的一组规范。还有第三种用法是，用一个实际的对

象作为一个榜样(exemplar),当作是指定类型的一个榜样对象。下面来自的乔伊斯(Joyce)《尤利西斯》(*Ulysses*)(1934,第183页)中的引文很好地阐明了这一点。

> ——斯蒂芬(Stephen)很客气地说,学者首先是男学生。亚里士多德曾经是柏拉图的学生。
>
> ——约翰·伊格林顿(John Eglinton)镇静地说,人们应该希望一直保持这样。人们会看见他,一位腋下夹着他的文凭的 *model* 男学生。

在绝大多数科学语境中,"*model*"很少在榜样的意义上使用。但是,在从全部细节中抽象出设计规范的意义上,我们接近于这种似乎是占有优势的用法。如何以更形式的和更数学的方式来说明这一点,是下一节的主题。不管怎样,在转向许多科学语境的引文时,我首先从形式的和数学的引文开始。[2]

18

> 满足一个理论 T 的所有有效语句的一种可能实现被称为是 T 的一个 *model*。
>
> (Tarski, 1953: p. 11)

引自物理学

> 在光谱学和原子结构领域内,对经典物理学的相似违背发生了。现有的大量证据表明,原子是由一个重的、带正电的核和绕核旋转的带负电的、像粒子一样的电子组成的。按照电荷之间相互吸引的库仑定律,电子只有绕核旋转,这样一个系统才不会立刻塌缩。但是,一个旋转的电荷,由于

---

① 详细分析可参见,Suppes, 1960a。

② 在这一节的所有引文中,我一直把"model"写成斜体字。

其加速度，将会发射辐射。因而同时提供了发光机制。

然而，这种机制与实验数据完全不符。很容易看到的两个主要困难是，第一，电子在原子中连续地旋转，原子应该一直发光。可是，在实验中，原子只在特殊的"受激"条件下才产生辐射。第二，根据这种 *model* 不可能说明单频（更正确地说，是一个窄范围的频率）谱线的发生。我们的 *model* 辐射电子会损失能量；结果，它不再能够保持自身与核之间的初始距离，而是会落向吸引的中心，当它落向中心时，会改变它的旋转频率。它的轨道将会螺旋式地终止在这个核上。根据电动力学理论，一个旋转电荷发射辐射的频率等同于旋转的频率，既然后者发生变化，前者就也应该发生变化。这样，我们的 *model* 不能说明谱线的清晰度。

(Lindsay and Margenau, 1963: pp. 390 – 391)

作者（吉布斯）认为，他的任务不是直接建立物理学理论，而是建构与热力学和其他物理学具有某些相似之处的统计力学的 *models* 之一；因此，他毫不犹豫地引入了统计特征的某些很特殊的假设。

(Khinchin, 1949: p. 4)

与德布罗意（de Broglie）、薛定谔（Schrödinger）和狄拉克（Dirac）的名字相联系的现代量子论（当然它用连续函数进行运算），借助于首先由马克斯·玻恩（Max Born）明确提出的一个大胆解释克服了这种困难：——方程中出现的空间函数不要求是原子对象的一个数学 *model*。这些函数只应该以数学方式来确定，如果我们进行一次测量的话，那些对象在一个特殊空间或一个特殊的运动态所出现的概率有多大。这个概念在逻辑上是无懈可击的，也取得了重要

的成功。但是,不幸的是,它迫使我们使用了维度数,不是从前物理学中的四维,而是随着构成所考虑系统的粒子数的增加而无限制地增加的一个连续统。我不得不承认,我自己觉得这种解释只是暂时有意义。我还认为,有可能给出实在的一个 *model*,也就是说,一个理论,将描绘事件本身,不只是描绘事件发生的概率。另一方面,似乎对我来说,在一个理论的 *model* 中,我们无疑不得不放弃这些粒子是绝对定域的概念。对我而言,这似乎是海森堡(Heisenberg)的不确定关系的正确的理论解释。可是,一个理论可以很好地存在,这个理论是真正意义上的一个原子理论(并不只是在特殊解释的基础上),在这个原子理论里,这些粒子在一个数学 *model* 中是没有定位的。

19

(Einstein, 1934: pp. 168 - 169)

引自生物学

一位化学家有两种重要方式来了解物质的三维结构。其一是运用取决于物质属性的物理学工具,来提供有关原子在分子中的相对空间位置的信息。为了达到这个目的所开发的最强有力的技术是 X 光的衍射。另一条进路是建立 *model*。正在考虑的分子的一个 *model*,是根据这些原子之间的键角和键距的精确值,运用原子出现于其中的比例 *model* 来建构的。这些参量的基本信息当然来自物理测量,比如,晶体研究中的 X 光的衍射,是晶体学中的一种常见技术。因此,我们对 DNA 的实际结构的理解和如何将其公布于众,必须一开始就考虑 X 光及它们如何与物质进行相互作用。

(Portugal and Cohen, 1977: p. 204)

为了开始研究这种简单网络的基因座的复制和色散效应，我完全忽略了复合、倒位、中间缺失、易位和点的突变问题，只运用转化来 *modelled* 色散。我使用了这样的简单程序：对于单倍染色体组来说，这个简单程序随机地决定，复制或转化是否在每一次重复中都会发生，在多大的基因座范围内发生，在哪两个基因座之间发生复制，以及转座到哪个位置。

即使作出了这些极大的简化，这个系统的动力学也是复杂的和缺乏研究的。既然简化 *model* 不会破坏基因座，一个基因座的形成率取决于现有的复制数，而且，预计每一个基因座都是近似随机地指数增长。这通过复制的空间范围——它提供对邻近基因座的正相关复制——作出修改，然后，再通过复制与转化的频率比作出进一步的修改。基因座的一阶破坏的进一步假设会降低指数增长率。然而，这里不进一步讨论这种动力学，因为这种简单 *model* 的主要目的是，在复制和转化的许多事例发生之后，考察这种调控性的体系结构。

(Kaufmann, 1982: p. 28)

引自社会科学

因此，经过对双方进行的明智比较而建立起来的理性选择的 *model*，似乎不很符合所描述的博彩情境中的理性行为的情况。

(Arrow, 1951: p. 21)

在建构 *model* 时，我们将假设，每一个变量都是小组成员的某种平均或合计。例如，测量 D 的方法是，在一把尺子上定位小组成员的意见，在位置刻度上标出数字，然

后，根据这些数字来计算成员意见的标准偏差。甚至能够把中间变量（尽管不进行直接测量）看成是个别成员的平均值。

(Simon, 1957: p. 116)

关于学习的数学 *model* 的这项工作，不是试图使任何一个特殊的行为理论体系形式化；不过，格思里（Guthrie）和赫尔（Hull）的影响是最值得关注的。与在数学理论化方面的较古老的企图相比，新近的工作更多地关注与这些 *model* 相关的数据的详细分析和直接检验这些 *model* 的定量预言的实验设计。

20

(Bush and Estes, 1959: p. 3)

引自数理统计

我将描述……在接受一个观察的随机过程的数学 *model* 时所用的各种不同标准……例如，根据时间 t，考虑规定地点的车通量。第一个假设是典型的数学假设，它是根据事实提出的，而且用来简化数学运算。该假设是，这个 *model* 的随机过程与增量无关……下一个假设（即固定增量假设）声明，如果 $s<t$，那么，$x(t)-x(s)$ 的分布只取决于时间间隔的时长 $t-s$。这个假设意味着，我们不可能让时间既指低谷期，也指高峰期。交通强度必须是恒定的。

下一个假设是，事件每次只发生一次。这个假设至少对一位数学家来说是很自然的。由于测量精度有限，所以它对于一位观察者来说毫无价值……下一个假设是一个更加定量的类型，这个假设对于已经明白了泰勒（Taylor）定理的任何一个人来说，也是自然的。在时长为 h 的时间间

隔内至少有一辆车通过的概率应该是 $ch+o(h)$。

(Doob，1960：p. 27)

引自应用数学

否定无穷概念，就像它不是美国人的一样，是非数学的。然而，恰好是在这种数学异端(mathematical heresy)的一种形式之上，离散的 *model* 理论建立起来了。

(Greenspan，1973：p. 1)

关于引文的评论。首先，第一段引文摘自数理逻辑，下面的三段引文摘自物理学，接下来的两段引文摘自生物学，再下面的三段引文摘自社会科学，接着的两段引自数理统计，最后一段引文摘自应用数学。这些引文无论如何没有穷尽容易收集到的各种不同用法。

完全有理由认为，不可能把这些引文中出现的 *model* 这个词的几种用法归于一个概念。我认为，声称在所有这些引文中只在精确地相同的意义上使用 *model* 这个词太言过其实。来自杜布(Doob)的引文体现了一种很普遍的趋势，即混淆或合并逻辑学家所谓的模型和模型的理论。用 *model* 这个词指理论的定量假设的集合，在数理统计中和在行为科学中是很普遍的实践。也就是说，这个句子集合，经过精确处理后会被当作是公理，或者，假如它们本身还不太明显，将构成阐述一个公理集的直觉基础。在这种用法中，一个模型是一个语言实体，而且与塔斯基(Tarski)的定义描述的用法截然不同，根据塔斯基的定义，一个模型是符合一个理论的非语言实体。

我们也应该注意到，*model* 这个词在计量经济学中的一种特定技术用法。在这种意义上，一个模型是逻辑学家意义上的模型类，并且逻辑学家所谓的一个模型，被计量经济学家称为一

种结构。

对我来说,这些似乎都不是严重的困难。我主张,如果没有
21 曲解的话,塔斯基意义上的模型概念可以被用作摘取上述引文的所有学科中的基本概念。在这种意义上,我会断言,模型概念(的含义)与在数学和经验科学中的含义是一样的。这些学科之间的差别体现在它们对这个概念的用法中。当在含义的恒定性和不同用法之间作出这种比较时,人们不参考一个概念的用法如何来说明它的含义,这个困难的语义问题其实有时并不会发生。当我说到一个模型概念的含义时,我将总是在一个明确定义的技术语境中来说的。我的主张是,给定一个模型概念的这种技术含义,数学家质问关于模型的这一类问题,而经验科学家往往质问关于模型的那一类问题。

通过分析 *model* 这个词在上述引文中的用法来为 *model* 概念的这个论点作出辩护是有启发性的。正如已经表明的那样,摘自塔斯基的引文代表了 *model* 在数理逻辑中的一个标准定义。在这一点上,我将不考虑对模型的技术描述,这将在本章的下一节进行。大体上说,一个理论的可能实现是一个适当逻辑类型的集合论实体。例如,群论的一种可能实现是任何一个有序对,它的第一项是非空集,第二项是这个集合上的一种二元运算。因此,一个模型恰好是符合该理论的一种可能实现。我们需要的一个重要区别是,一个理论是由一个句子集合组成的一个语言实体,而模型一般说来是符合该理论的非语言实体。(理论能够被看成是命题的集合体,而不是句子的集合体。这并不影响这里提出的主要论点,但是它会改变下一节的进路。)

我认为,模型概念在林赛(Lindsay)和马尔诺(Margenau)的引文中的用法,能够以下列方式用这些术语作出重新讨论。把原子轨道理论阐述成为一个理论。那么,产生的问题是,按照

与实验密切相关所定义的实体，这种理论的一种可能实现，究竟是否实际上能被看成是该理论的一个模型；或者，以更简单的方式说，一个轨道理论的模型与从原子现象的物理实验中获得的数据完全相符吗？许多物理学家希望把原子的轨道理论模型不止看成为一种特定的集合论实体，这是事实。他们把它设想成是建立在类似太阳系基础上的一个很具体的物质。我认为，重要的是指出，这两种观点实际上并不矛盾。从形式上把一个模型定义为一个集合论实体，即它是由对象、关系和关于这些对象的运算的一个集合构成的特定类型的有序元组，并没有排除吸引物理学家的这类物理模型。物理模型可以被简单地用来定义集合论模型中的对象集合。由于这种观点是重要的，对此作出更加详细的阐述是合适的。我选择以第七章详细地讨论的经典质点力学为例。我们可以根据一个质点集合 P 的五个基本概念把经典质点力学公理化，与运行时间相符合的实数间隔 T、根据质点集合和时间间隔的笛卡儿积定义的位置函数 s、根据质点集合定义的质量函数 m 以及根据质点集合、时间间隔和正整数集合的笛卡儿积定义的力函数 f（正整数集合参与力函数的定义只是为了提供命名力的方法）。经典质点力学公理的一种可能实现就是一个有序的五元组 (P, T, s, m, f)。经典质点力学的一个模型将是这样一个有序的五元组。（事实上，第七章提供了一种更复杂的分析。）很简单地可以看出，在物理学家的经典质点力学意义上的一个实际的物理模型，如何与这种模型的集合论意义相关。例如，在太阳系的情形中，我们只能把质点集合看成是星体的集合。另一种稍微更加抽象的可能性是，把质点集合看成是星体质心的集合。这通常例示了这种情况。一个理论的一种集合论模型含有这样一个基本集合：它是由通常认为是构成所考虑的物理模型的对象所组成的，而且可以用同样

22

的方式来考虑给定的原始关系和函数。我们在前面的段落里已经用到短语“集合论模型”和“物理模型”。争论 *Model* 这个词的哪一种用法在经验意义上更基本或更适当,似乎是无用的。我自己的论点是,集合论的用法是更基本的。很有物理学思想或经验头脑的科学家(他们可以不同意这种论点,并认为在给定的经验科学分支中,物理模型是更重要的)可能更同意我的系统评论。

关于这种观点的一个历史事例是,凯尔文(Kelvin)和麦克斯韦(Maxwell)为了寻找电磁现象的力学模型所付出的努力。[①]毫无疑问,他们两人是在字面的物理意义上考虑可能的模型,但是,很容易把他们关于这个论题的已出版的论文集,重新看成是寻找说明所观察到的电磁现象的连续介质力学理论的集合论模型。此外,事实上,在他们的论文集中,正是形式的部分具有永久的价值。从根本上说,正是麦克斯韦的数学理论,而不是像弹性固体那样运转的以太的物理图像,证明对下一代科学家是永远有用的。无论如何,在科学的许多方面,历史的甚至是当前的图像模型的重要性是显而易见的(参考,Harré, 1999)。

第三段引文摘自欣钦(Khinchin)关于统计力学的著作。“作者”这个短语是指吉布斯,欣钦此时是在讨论吉布斯的工作。在这段引文中,*Model* 这个词的用法特别地支持了集合论的观点,因为欣钦主张,吉布斯在他关于统计力学的基础的工作中,并没有直接地求助于物理实在或试图建立真的物理理论;更确切地说,他试图建构部分地与热力学的复杂的经验事实和物理学的其他分支类似的模型或理论。也许,比杜布的情形还要更加直接,这段引文例示了把理论与模型的逻辑类型混合在一起

① 作为对自然现象展开数学分析的一个例子,这些努力是勇敢的。参见,Kelvin (1846) and Maxwell (1861 - 1862)。

的趋向;但这也不会造成困难。吉布斯和欣钦的工作都很容易进行直接的阐述,以便明显地和准确地承认,在数理逻辑中理论与模型之间的区别。吉布斯的统计力学著作的抽象性,为应用逻辑学家所用的模型的精确概念提供了一个极好的事例,因为根本没有直接的趋向,把吉布斯的统计力学模型看成是一个物理世界的**那种**理论。

23

物理学中的 *model*。我认为,就物理学中使用 *model* 这个词来说,下列观察在经验上是可靠的。在从前已确立的与其试图说明的经验现象非常相符的物理学分支中,曾经运用 *model* 这个词的趋向很不明显。理论语言、实验语言和常识语言混合在一起构成了一个实际的整体。理论语句被断言为好像它们是描述世界的一种方式。实验结果被描述为好像只有一种明显的语言描述它们。常识的概念(也许到处被提炼)被适当地看成与物理学理论是同类的。另一方面,在迄今为止对其所关注的详细的物理现象提供的说明是不充分的物理学的那些分支中,它们更加频繁地使用了 *model* 这个词。使用这个词的内涵是,这种模型像一架飞机或一艘船的模型一样。这彻底地简化了真正的物理现象,而且只说明某些主要方面。此外,在这个词的这些用法中,将要强调的是,在作为物理对象或非语言对象的模型与作为一个理论的模型之间,始终有相互影响。摘自爱因斯坦的引文以一种有趣的否定方式举例说明了这一点,爱因斯坦对量子力学中缺乏粒子运动的适当的数学模型而深感悲哀。

生物学中的 *model*。在对 *model* 这个词的使用上,摘自生物学的两段引文,与作为一种明确地简化复杂的实际现象的物理模型概念没有什么显著区别。似乎很可能生物实体的结构的复杂性必定会在不确定的未来保持正在流行的这种简化。[假使引自生物学的第二段引文(即考夫曼(Kauffman)的引文)的

语境有任何可能的不确定性的话,它直接与现代遗传学理论相关。]

社会科学中的*model*。摘自阿罗(Arrow)的引文例示了在社会科学中的一种同样必要的趋势,即遵循极端简化的过程。阿罗谈到了理性选择的*model*,因为他想到的这个理论并没有对其所关注的现象给出充分的描述,而是只提供了一个高度简化的框架。同样的评论也相当好地适用于摘自西蒙(Simon)的引文。在西蒙的引文中,我们拥有在社会科学和行为科学中很普遍的一种例示的附加现象。科学家用广泛而普遍的术语陈述一个特定的理论。为了检验这个理论,完成某些定性的实验。由于这些实验的成功,喜欢更加量化的和更精确的理论的科学家,致力于对最初的理论进行所谓的“模型建构”。在逻辑学家的语言中,更加适当的说法是,与其说是在建构一个模型,不如说他们的兴趣是在建构与最初理论的直觉想法相匹配的一个定量理论。

24 在摘自布什(Bush)和埃斯蒂斯(Estes)的引文和摘自杜布的引文中提出了一个重要的思路,这种思路事实上与逻辑学家使用的模型概念密切相关。我这里考虑数理统计中的模型概念、有关评价模型中的参量的大量文献,以及检验这些参量的假设。通常情况下,在评价一个模型的参量的统计讨论中,把这种讨论转化为术语的用法在其中与逻辑学家的用法完全一致的讨论,通常是一项无价值的任务。详细考虑统计问题几乎要求把模型考虑为是数学的或集合论的,而不是简单的物理实体。对于统计学家和行为科学家而言,“模型如何很好地与数据相符合”的问题是一个很自然的问题。只是最近物理学家才开始这样认为,并且确实,他们还用具有统计学基本风格的术语来报告物理学中的某些实验工作。

在引用和分析这些引文时，我设法论证，数理逻辑学家所用的模型概念是经验科学的任何一门分支的精确陈述所需要的基本概念。为了认同这种论点，没有必要排除或悲叹经验科学中现在盛行的模型的不同用法或不同概念。正如已经表明的那样，我自己准备承认在物理学和工程领域当前讨论很多的物理模型概念的意义和实践重要性。我曾试图主张，在理论的精确陈述方面或在数据的精确分析方面，逻辑学家意义上的模型概念就是这种恰当的概念。

本书其余的绝大部分所关心的，是表明如何可以用模型的集合论概念来澄清和加深科学哲学中对一些一般论题的讨论。然而，本章的后面部分仍然主要是澄清模型概念本身。迄今为止所讨论的内容还不是很准确或详细。我现在转向各种更加形式的考虑。

2.2 具有标准形式化的理论

一个具有标准形式化的理论是在具有恒等式的一阶谓词逻辑中阐述的理论。一阶逻辑的常用逻辑工具是假定的，主要变量的范围包括元素和逻辑常项的相同集合，特别是句子的连接词和通用的存在判断量词以及恒等号。存在三个非逻辑常项：谓词或关系符号、运算符号和个别常项。

这种理论的表达，即这种理论的有穷序列的语言符号，被划分为术语和公式。通常给定每一个递归定义，然后我们直接考虑几个简单的事例。最简单的术语是变量或个别常项。新的术语是通过以适当方式把较简单的术语与运算符号结合起来建立的。原子公式是由一个谓词和适当数量的术语组成的。分子公式，即组合公式，是借助于句子连接词和量词由原子公式构成

的。从逻辑教科书中对这些问题的任何一种讨论来看，对如何做到这一切的一般描述是相当熟悉的。我们在这里不考虑细节。通过举例可以很容易、很好地表达这种直观的想法。在考虑这些事例时，我们不会过多地讨论关于一阶逻辑的明确本性，而是假定已经对此相当熟悉。在一个具有标准形式化的理论中，首先，我们必须从术语的递归定义和以该理论的原始的非逻辑常项为基础的公式开始。第二，我们必须说出把该理论的哪些公式看成是公理。[①]（适当的逻辑推理规则和逻辑公理不作进一步讨论而被假定为是有效的。）

25

事例：定序测量。作为具有标准形式化的理论的第一个事例，我们可以考虑定序测量(ordinal measurement)的简单理论。首先，我们使用标准的逻辑符号。句子的连词“&”、“∨”、“→”、“↔”和“¬”分别具有“与”、“或”、“如果……那么”、“当且仅当”和“非”的含义。我用“∀”表示全称，“∃”表示存在量词。对于标点符号而言，以常见的方式使用括号，当实践中不会产生歧义时，也去掉括号。提到等号“＝”和变量“x”、“y”、“z”……结束了对逻辑符号的描述。

这里作为一个例子考虑的定序测量理论 $\mathfrak{O}$ 的语言含有一个原始的非逻辑常项：二元关系符号“≥”。因为 $\mathfrak{O}$ 的语言在其原始的非逻辑常项中没有运算符号，所以，$\mathfrak{O}$ 的术语的定义是极其简单的。[②] $\mathfrak{O}$ 的术语恰好是变量“x”、“y”、“z”……我们用这

① 更为需要说明的是，我们如何描述或挑选一个理论的公理，还有，该理论的一个有效语句的概念如何可以为了达到某些目的取代一个通常的公理化理论的概念，但我认为，在这里能够安全地忽略掉这些问题。关于一种极其直观的讨论，读者可参考塔斯基著作的头几页(1953)。

② 为了简化说明，在这一段和后面，我说“$\mathfrak{O}$ 的术语”，而不是“$\mathfrak{O}$ 的语言的术语”。在这一节的语境中，用“$\mathfrak{O}$”既指语言，又指理论，应该不会造成混乱。

些变量组成 $\mathfrak{D}$ 的公式。首先,我们定义 $\mathfrak{D}$ 的原子公式。$\mathfrak{D}$ 的一个原子公式是三个符号的一个序列,即,一个变量,接着是"="或"≥",再接着是一个变量。这样,"$x=y$" 和" $z \geq z$ "是 $\mathfrak{D}$ 的原子公式的例子。根据原子公式,我们就可以给出所谓公式的一个递归定义。

(a) 每一个原子公式都是一个公式。

(b) 如果 S 是一个公式,那么$\neg(S)$也是一个公式。

(c) 如果 R 和 S 都是公式,那么$(R)\&(S)$、$(R)\vee(S)$、$(R)\rightarrow(S)$和$(R)\leftrightarrow(S)$也都是公式。

(d) 如果 R 是一个公式,v 是任何一个变量,那么$(\forall v)(R)$和$(\exists v)(R)$都是公式。

(e) 一个表达式,只有在遵守上面的规则时,才是一个公式。

公式的这个定义之所以被称为递归的,是因为给定 $\mathfrak{D}$ 语言的任何一个有穷序列的符号,我们都能够在有限的步骤内用这个定义来确定这个序列是否是 $\mathfrak{D}$ 的一个公式。同样应该显而易见的是,对于所有具有标准形式化的理论来说,这个定义恰好也是一样的。从理论到理论所发生的变化是术语和原子公式的定义。

$\mathfrak{D}$ 的两个公理恰好是两个句子: 26

1. $(\forall x)(\forall y)(\forall z)(x \geq y \ \& \ y \geq z \rightarrow x \geq z)$

2. $(\forall x)(\forall y)(x \geq y \ \vee \ y \geq x)$

直观地看,第一个公理是说,≥所指定的关系是可传递的,第二个公理是说,它是一种弱连接[关于这些问题和类似的关系特性的讨论,可参见 Suppes(1957/1999,第十章和 1960b,第三章)]。

下面我转向该理论的模型定义。首先,我们定义 $\mathfrak{D}$ 的可能实现,这在 2.1 节已经提到过。设 A 是一个非空集,≥是在 A

上定义的一种二元关系,即$\geq$是笛卡儿积$A \times A$的一个子集。于是,结构$\mathfrak{A}=(A,\ \geq)$是$\mathfrak{O}$的一种**可能实现**。有人可能注意到,集合A被称为实现$\mathfrak{A}$的**定义域或全域**。我会假定在什么情况下$\mathfrak{O}$的一个句子在直观上是明确的,即没有自由变量的$\mathfrak{O}$的一个公式是一种为真的可能实现。[在这个意义上关于真理的一个明确定义可参见塔斯基的经典著作(1935)。]我们于是说,$\mathfrak{O}$的一个*model*是$\mathfrak{O}$的公理为真的$\mathfrak{O}$的一种可能实现。例如,设

$$A = \{1,\ 2\}$$

和

$$\geq = \{(1,\ 1),(2,\ 2),(1,\ 2)\}$$

那么$\mathfrak{A}=(A,\ \geq)$是$\mathfrak{O}$的一种可能实现,恰好是建立在它的集合论结构的基础之上的,但也容易核对,$\mathfrak{A}$也是$\mathfrak{O}$的一个模型,因为关系$\geq$在集合A中是可传递的和弱连接的。另一方面,设

$$A' = \{1,\ 2,\ 3\}$$

和

$$\geq' = \{(1,\ 2),(2,\ 3)\}$$

那么$\mathfrak{A}'=(A',\ \geq')$是$\mathfrak{O}$的一种可能实现,但是,它不是$\mathfrak{O}$的一个模型,因为关系$\geq'$在集合$A'$中既不是可传递的,也不是弱连接的。

以公理方式建立的理论。本章,我从塔斯基的定义开始,他把一个理论的模型定义为,满足该理论的所有有效句子的任何一种可能实现。就在上面,我通过在这种可能实现中该理论的公理都得到满足的要求,陈述了这种表述的另一种形式。当把一个理论的有效句子集合定义为是该理论的给定的公理集及其

逻辑结论时，该理论就被说成是*以公理化方式建立的*。一般来说，本书只限于分析以公理化方式建立的理论，但要在下一节而非本节所确立的框架之内。正如我们将要看到的，这不是对具有明确科学内容的任何一个理论的严格约束。

于是，假如承认定序测量理论模型恰好是满足该理论的两个公理的可能实现，我们可能会质问关于理论 $\mathfrak{D}$ 的另外一类问题。

我们要问的这类问题自然属于特定的类。一类问题是关于该理论的模型之间的关系问题。在数学中，对这种结构问题的研究是普遍的，而且正如我们会看到的，这种研究不要求为理论 $\mathfrak{D}$ 给出一种标准的形式化。另一类问题是关于这种形式化的问题。例如，可判定性的元数学问题——存在着一种机制决定程序来判断该理论的一个公式是不是该理论的一个有效句子吗？ 27

第三类问题是关于该理论的经验解释和检验问题，本书不准备详细地讨论这类问题，但必须一提的是，这类讨论几乎从来不需要一个理论的一种标准形式化。

科学形式化的困难。我已经写道，好像给出一个理论的标准形式化的决定，完全依赖于形式化所伴随的有用性问题。不幸的是，这种决定并非这么简单。我想提出的一种观点是，在经验科学中，大多数理论的一种简单的标准形式化是不可能的。这种困难的来源很容易描述出来。几乎所有有任何趣味或力量的系统化的科学理论都把大量的数学假定为是它们的形式背景的一个组成部分。根本没有一种简单的或一流的方式，在只假定初等逻辑方法的一种标准形式化中包括这种数学背景。单就这一点导致了科学哲学家对科学理论结构的大多数讨论与对这些理论的标准的科学讨论之间缺乏联系。[对这种观点的支持，

参见 van Fraassen (1980)。]

因为所讨论的这种论点为采纳下一节概述的集合论进路提供了一个重要论据，所以较详细地考虑一两个事例将是合适的。

假设我们希望给出初等概率理论的一个标准形式化。一方面，如果我们遵循这种标准的进路，我们就需要有关集合的公理来理解关于联合发生的两个事件的讨论，因为这两个事件被看成是集合，联合发生被看成是事件的集合论的交集。另一方面，我们也需要关于实数的公理，因为我们希望谈论事件的数值概率。最后，在陈述了关于集合的一组公理和关于实数的另一组公理之后，我们能够做的是，像通常设想的那样，陈述恰好属于概率理论的公理。在这一堆混乱的公理中，人们很容易会忽略掉那些特殊的有关概率的公理。更重要的是，每当我们考虑把一个科学理论形式化时，不断地复述关于集合和关于数的这些一般公理，是无意义的和无趣的。没有人有充分的理由来这样做。

第二个重要的例子是力学。在这种情形中，我们需要用实数的基本理论和许多标准的数学分析，特别是与解微分方程相关的那类分析，来阐述引人感兴趣的问题。再者，作为一种绝技，我们能够用几条进路中的任何一条进路，把附加的许多数学概念带入力学的一种标准形式化当中，但这种结果太不实用。

不幸的是，因为一个科学理论的标准形式化通常是不可行的，所以，某些科学哲学家和科学家似乎认为，任何一条严密的进路都是不可能的。下一节的主要目标是表明，事实并非如此。

28

有用的形式化事例。① 在转向这项任务之前，我想考虑这

---

① 在不失任何连续性的前提下，可以删掉相当技术性的这一小节。

样一个例子：根据理论的标准形式化对理论进行的处理，如何能够拥有令人感兴趣的科学哲学结论。我想到的这个问题是特殊测量理论的有穷可公理化之一。为了准备回答这个问题，我们可以暂时回到定序理论 $\mathfrak{D}$。我们这里不必证明直观上显而易见的下列事实：对于理论 $\mathfrak{D}$ 的任何一个有穷模型 $\mathfrak{A}=(A, \succeq)$（也就是说，集合 A 是有穷的）而言，都存在着一个在 A 上定义的一个实值函数 f，使得对于 A 中的每一个 x 和 y，

$$x \succeq y\text{，当且仅当，}f(x) \geqslant f(y) \qquad (1)$$

用代数的术语来说，$\mathfrak{D}$ 的任何一个有穷模型都与数值模型是同态的(homomorphic)，正是这个简单的定理成为把理论 $\mathfrak{D}$ 称为一个测量理论的基础。（下面更加详细地讨论的一种观点是，对于 $\mathfrak{D}$ 的任意模型来说，这个定理当然不是真的。）

对于在心理学中的许多应用而言，从定序理论过渡到下列理论是很自然的，在这个理论中，对象的排序位置之差也是有序的。因此，我们会质问，必须把哪些公理强加于一个四元关系 D 和 D 在其上被定义的集合 A，使得存在着一个具有下列特性的实数值函数 f：对于 A 中的每一个 x、y、z 和 w 来说，[1]

$$xyDzw\text{，当且仅当，}f(x)-f(y) \geqslant f(z)-f(w) \qquad (2)$$

（关于把四元差关系 D 应用于心理学的偏好理论和测量理论的广泛讨论，可参见 Suppes and Winet 1955；Davidson, Suppes and Siegel 1957; Luce and Suppes 1965; Suppes and Zinnes 1963; and Krantz et al. 1971。）我们能够用等式(2)获得这种差关系理论的一个直接的外在公理化，这种差关系理论很自然地被称为*超定序*测量(hyperordinal measurement)理论。正如在

---

① 我在第三章用一个通常的弱的不等号“$\geqslant$”来取代“D”。

第一章中所讨论的，这当然恰好是使定义明确的问题不可能成为内在公理化的一种方式。在超定序测量理论 $\mathcal{H}$ 的标准形式化的技术意义上，我们需要基本公理，使得理论 $\mathcal{H}$ 的一种有穷实现 $\mathfrak{A}=(A, D)$ 是 $\mathcal{H}$ 的一个模型，当且仅当，存在着一个在 A 上的实值函数 f 使得(2)成立。还有必要强加其他限制，使这个问题令人感兴趣。如果我们愿意承认，这些基本公理都是可数的无穷，那么，通过为每一个 n 列举所有可能的同构类型，可能会找到一个肯定的正解，但是，这类公理化很少能深入洞察到理论 $\mathcal{H}$ 的任何一个模型都必须拥有的一般的结构特征。因此，我们转向一种有穷公理化的可能性。不幸的是，有穷可公理化的问题是很难解决的。然而，在当前的情形中，我们希望该理论的所有有穷模型都满足我们的公理；这样，要求该理论的所有公理都是全称句是合理的，即把句子的所有量词放在句子的前面，句子其余部分作为它们的范围。这种观点是，这样的句子在子模型下是闭合的。这意味着，如果一个句子适合于一种实现 $\mathfrak{A}=(A,D)$，那么在 B 是 A 的一个子集和 D' 是只限于集合 B 的关系 D 的地方，它也适合于 $B=(B, D')$。在子模型下的这种闭合是全称句特有的，如果我们打算将该理论的所有有穷模型都公理化，那么这样一种闭合特性也是需要的。

29

正如已经证明的那样，理论 $\mathcal{H}$ 的结构当用基本的全称句来表示时是相当复杂的，因为达纳·斯科特(Dana Scott)和我(1958)能证明，全称句的有穷集合没有一个能为 $\mathcal{H}$ 提供一个适当的公理集。证明方法是，随着不断增大的定义域和根据具有下列特性的每一种实现来构造 $\mathcal{H}$ 的一个可能实现的有穷序列：在这种实现的定义域内删掉任何一个元素都会产生 $\mathcal{H}$ 的一个模型，但是完全实现本身并不是这样一个模型。从沃特(Vaught, 1954)的一般公理化标准得出，当能够找到这样一个

实现序列时，该理论不可能通过全称句的一个有穷集得到公理化，或者，通过一个全称句进行公理化，结果也是一样的。[①]

这类否定结果特别取决于在标准形式化框架内运行。事实上，对于理论在标准形式化框架之外的公理化而言，如果有任何可使用标准的话，也是很少见的。对于像 $\mathcal{H}$ 那样的一个理论来说，刚才描述的那类否定结果更有意义地说明了它的结构。但是由于前面陈述的理由，很少有科学理论具有理论 $\mathfrak{D}$ 和 $\mathcal{H}$ 的简单性。事实上，很难在测量领域之外找到具有可比较简单性的其他有经验意义的理论事例。

斯科特(1964)依据支持基本公理的一个无穷集的框架给出了对 $\mathcal{H}$ 的一种极好的公理化。[②] 然而，他的框架是根据 $\mathcal{H}$ 的模型域的元素的所有可能排列来表示的，这种框架不会为我们从像传递性(transitivity)或连接性(connectivity)那样的公理获得的理论结构提供那种内在感。此外，关于根据全称句的有穷公理化的否定结果表明，寻找 $\mathcal{H}$ 的一种简单的、直观的内在描述可能是无望的。

一旦我们试图既获得无穷模型，又获得有穷模型，那么，在一阶谓词逻辑中，即在具有作为惟一非逻辑基本常项 $\geq$ 的一种标准形式中，把定序理论 $\mathfrak{D}$ 描述为一种测量理论也是不可能的。这种困难的核心在于这样的事实：如果一个一阶公理有一个无穷模型，那么，它就有无界基数的模型，而且，从这个事实很

---

① 泰特(Tait, 1959)已经表明，在确定关于有穷模型的有穷可公理化的一般结果时，即使要求子模型下的闭合，只考虑全称句也是不可能的。如果不局限于全称句，关于有穷公理化性的更一般的结果就会复杂得这里无法讨论。关于林德斯特伦(Lindstrom)对 $\mathcal{H}$ 不是有穷可公理化的先前未公开的结果的提出，可参见 Luce 等人(1990)的文献。

② 斯科特关于概率的一个相关结果作为 5.7 节的定理 3 来陈述。

容易断定，对于我们强加的任何一个附加公理的无穷模型来说，我们不可能证明存在着一个满足(1)的真值函数 f。一个不太复杂的问题来源于这样的事实：一个任意大的基数模型仍然可能与一个数值模型是同态的，但是，这能够通过假设$\geq$是一种反对称关系，即通过补充公理：

30

$$(\forall x)(\forall y)(x \geq y \ \& \ y \geq x \rightarrow x = y)$$

大约相当直接地获得和首先避免。一旦包括了这种公理，任何一种同态一定也是一种同构，当然，不可能为一个大于闭联集基数的基数模型找到任何一个同构的数值模型。为了保证 $\mathfrak{D}$ 的无穷模型与数值模型是同态的，强加于 $\mathfrak{D}$ 的无穷模型的充分必要条件并不是一阶条件，即不是用理论 $\mathfrak{D}$ 的语言来阐述的，但是，它确实在下一节讨论的集合论框架内有一种自然的表达，第三章明确地考虑这个问题。

2.3 通过集合论谓词定义的理论[①]

尽管由于上一节提供的那些理由，有经验意义的大多数科学理论都能作出标准的形式化这种承诺并不是切实可行的，但是，有一条进路相当精确并满足现代数学的所有严格标准，能对这些理论进行一种公理的形式化。从形式的观点来看，这条进路的本质是把集合论的公理增加到初等逻辑的框架内，然后，在这种集合论的框架内使科学理论公理化。从本书其余部分关于

---

① 这一节的一般观点与苏佩斯(Suppes)在许多年前的最后一章(1957/1999)中阐述的观点是一样的。关于这种观点和相关观点的更一般讨论，可参见 Suppes (1974)和 van Fraassen (1980)。

这个论题的观点来看,我们为一种公理化基础挑选什么样的特别不同的集合论这点并不重要。从一种操作的观点来看,最好能把我们的所作所为描述为是在朴素集合论中进行的运算。公理化集合论和朴素集合论之间一个重要的区别是,在公理化集合论中,人们必须不断地考虑关于集合的存在问题,因为这恰好是会产生悖论和不一致性的那些问题。在朴素集合论中,人们会一直进行下去,好像根本不存在任何悖论,而且,无忧无虑地假设,所有的集合都以一种预期的直觉方式存在着。尽管在这里将不明显地考虑集合的存在问题,但将本书接下来所阐述的一切嵌入某种明确的合理化集合论之中是一件简单的和平常的事情,例如,在包括我自己在内的几本书中(Suppes 1960b)详细地提出的和描述的策梅洛-弗伦克尔(Zermelo-Fraenkel)的看法。关于这种观点的保留意见,可参见本节的最后一部分。

许多年来,我的口号一直是"使一个理论公理化就是定义一个集合论谓词"。或许关于这个口号需要澄清的最重要的混淆之处是公理化和定义之间的密切关系。我们可能会先考虑一些公理化理论的纯数学例子。正如我们在最后一节所看到的,我们能够通过为理论提供一种形式化,把它们简单地看成是完全自主的学科,在系统的意义上,这只需要初等逻辑,并不需要其他数学背景。然而,在当代数学家当中,这很少是做数学的标准方式,因为已经澄清的原因,即在这种限制的框架内处理任意复杂性的理论的棘手。

事例:群论。让我们首先考虑一个群(group)的公理,这现在通常在高中数学课本中有所讨论。下面是一个标准的形式: 31

A1. $x \circ (y \circ z) = (x \circ y) \circ z$

A2. $x \circ e = x$

A3. $x \circ x^{-1} = e$

这里的$\circ$是指这个群的二元运算，e 是指单位元素（identity element），$^{-1}$是指逆运算。孤立地理解的话，这些公理的困难是人们没有完全理解，如何把它们与其他理论的其他公理联系起来，或更准确地说，如何把它们与数学对象本身联系起来。这些不确定性很容易通过下列方式来搞清楚：即承认这些公理本质上是一个定义的组成部分，也就是说谓词的定义“是一个群”。这些公理是这个谓词定义的最重要的部分，因为它们告诉我们，满足谓词“是一个群”的数学对象，或换句话说，是一个群的一个数学对象，必须拥有这些最重要的特性。为了使这些公理成为确切定义的一部分，必须先解决一两个技术问题，而且，为解决这些问题提供方法是我们的集合论框架的主要功能之一。由这些公理可见，在某种意义上，一个群一定是具有一种二元运算、一个单位元素和一种逆运算的一个对象。我们如何精确地谈论这样一个对象呢？一种建议是，用下列方式阐述这种定义：

> 一个非空集 A 是关于二元运算$\circ$、单位元素 e 和逆运算$^{-1}$的一个群，当且仅当，A 中的每一个 x、y 和 z 都满足上面给出的三个公理。

关于这个定义要注意的第一个问题是，它不是一元谓词（one-place predicate）的定义，而实际上是一个四元谓词（four-place predicate）的定义。刚才阐述定义的方式在某种程度上掩盖了这一点。如果我们在集合论的范围内为它提供一个精确的阐述，那么，该定义会明确表明，不仅字母“A”，而且运算符号和单位元素的符号都是把值当作任意对象的变量。因此，沿着上述思路，只在集合论的范围内，一种形式上精确的定义如下：

> A 是关于$\circ$、e 和$^{-1}$的一个群，当且仅当，A 是一个非空集，$\circ$是在 A 上的一种二元运算，e 是 A 的一个元素，$^{-1}$是

> 在 A 上的一元运算，使得 A 中的每一个 x、y 和 z 都满足上面给出的三个公理。

后一种看法的实际困难是它很自然地把群作为对象来谈论。但是，在这种看法中，我们反而是谈论上面定义的一个四元谓词，即我们同时谈论一个集合和三个对象。我们下一步会讨论这种困难：

> $\mathfrak{A}$ 是一个群，当且仅当，存在着一个非空集 A、在 A 上的一种二元运算$\circ$、A 的一个元素 e 和在 A 上的一种逆运算$^{-1}$，使得 $\mathfrak{A}=(A, \circ, e, ^{-1})$ | 和 A 中每一个 x、y 和 z 都满足上面给出的三个公理。

这个定义的要点是使谓词"是一个群"成为一个一元谓词，因而提出把群作为明确的数学对象来谈论。正如从这个定义中能看到的那样，一个群是一个特定类型的有序的四元组。对一个群的这种集合论结构的描述回答了一个群是哪类对象的一般问题。在平常的数学实践中，定义所讨论的这类实体是很常见的，例如，有序的四元组，它在上一节定义的意义上对应于可能的实现，于是，在后面定义集合论结构本身时不再重复。例如，我们最好定义有序的四元组 $\mathfrak{A}=(A, \circ, e, ^{-1})$ 的代数，在这里，$\circ$是指 A 上的一种二元运算，e 是指 A 的一个元素，$^{-1}$ 是指 A 上的一元运算。引入了这个代数概念，我们就能把我们定义的群缩短为下列标准格式： 32

定义 1. 一个代数 $\mathfrak{A}=(A, \circ, e, ^{-1})$ 是一个群，当且仅当，A 中的每一个 x、y 和 z 都满足上面给出的三个公理。

最后这个定义所例示的格式是我们在本书接下来的部分最常用的格式。对于那些喜欢直接下定义的人来说，把刚才使用的陈述类型转化为在前面定义风格的意义上很明显的格式定义，是一件简单的和平常的事情。

为了使各种可能性更加明显，我们也能把代数看成是只由

一个非空集和在这个集合上的一种二元运算组成的结构。于是,把群稍微不同地定义为:

定义 2. 一个代数 $\mathfrak{A}=(A, \circ)$ 是一个群,当且仅当,对于 A 中的每一个 x、y 和 z,

公理 1. $x \circ (y \circ z) = (x \circ y) \circ z$ |

公理 2. 在 A 中存在着一个 w,使得

$$x = y \circ w \text{ |}$$

公理 3. 在 A 中存在着一个 w,使得

$$x = w \circ y$$

"集合论谓词"的含义。因为我已经在其他地方非常详细地阐明了定义的这些问题(Suppes 1957/1999, 第 8 章和 12 章),所以,我在这里除了作出一两个一般性的评论之外,不再对它们进行更深入的探究。首先,关于"使一个理论公理化就是定义一个集合论谓词"这个口号,最好再说些什么。"集合论谓词"这个短语是什么意思,可能还没有完全搞明白。这样的一个谓词只是能够以完全形式的方式在集合论中下定义的谓词。对于只基于属于关系(通常用符号"∈"表示)的基本谓词的一种集合论版本来说,这意味着,任何一个集合论的谓词最终都能根据属于关系单独地定义。在定义这样一个谓词时,我们将会自由地用到任意标准的数学概念,因为我们将假定,这些标准概念已经得到了全面发展和形式化。每当假定一个科学理论的上述集合论的形式化时,就如此明确地表明了这一点。事实上,本书的许多事例并不是这种情况。

33

集合论和各门学科。最后的评论是建议说,在一个蕴含纯数学概念的集合论定义和一个含有某一门特殊学科的集合论定义之间,可能存在着某种系统的差异。我并不打算提倡这种情

况。本书的论题之一是,在纯数学部分与理论科学部分之间作出明显区分的理论方式是没有的。举三个完全不同的例子来说,力学理论、热力学理论和学习理论的集合论定义,与纯数学的群论、环论、场论等的定义,是完全一致的。从哲学的观点来看,纯数学与应用数学之间根本没有明确的区分,尽管许多说法与此相反。这里,纯数学与应用数学之间或数学与科学之间的这种连续性,将会通过这两个领域内的许多事例来说明。

基本结构。这一部分的观点提供了本书的一般观点,我们能把这种观点确切地表述为用集合论的谓词来定义结构的类。这种结构的类是由满足该谓词的那些结构组成的。当这条进路是集合论的进路而不是逻辑学意义上的句法的或形式的进路时,所描述的这条进路类似于著名的布尔巴基(Bourbaki)的数学进路。(布尔巴基是一群数学家的笔名,这些数学家联合撰写了涵盖大部分数学领域的许多卷专著。)熟悉布尔巴基进路的人会认为本书迄今的论述是不能令人满意的,就我所描述的进路意在像关于数学的布尔巴基进路一样,但把这条进路应用于经验科学。这种批评的根源很容易说明。布尔巴基在许多地方,但尤其是在一般论文(1950)中断言,这根本不是在描述数学的各个不同部分的结构的意义上简单地给出独立的公理化问题。关于布尔巴基纲领更重要的是识别基本结构(布尔巴基称之为母结构)。已知这种识别,然后,集中于谨慎地探讨在数学的许多不同部分使用各种基本结构的方法。布尔巴基意义上的基本结构的简单例子是像前面定义的群、排序关系或序结构,还有拓扑结构。

还有另一点特别重要。在各个不同的地方,布尔巴基讨论了产生最熟悉的结构的方法。达·科斯塔(Da Costa)和丘瓦基(Chuaqui)(1988)已经根据涵盖了本书定义的很多结构的已知基本集合的笛卡儿积和幂集(power sets)给出了一种分析。在

这种情况下,这条进路将表明,在一个给定对象集上的几步基本运算如何足以产生有趣的结构。

似乎在我看来,达·科斯塔和丘瓦基生成结构的进路是一条富有成效的进路,而且,能适用于许多不同科学学科的理论研究。另一方面,对像布尔巴基认为对数学是重要的基本结构的这种识别抱有希望,似乎还为时过早。

34

与本节提出的思想框架相关,还有一种已经讨论过的结构主义的意义。这里,我指的是斯尼德(Sneed, 1971)、施特格米勒(Stegmüller, 1976, 1979)、穆利纳(Moulines, 1975, 1976),穆利纳和斯尼德(1979)以及巴尔泽尔(Balzer, 1978)的工作,还有这些作者的许多其他论文或著作。现在,这方面的文献很多,不可能作出详细的概述。一些重要的观点是由以斯尼德、施特格米勒和他们的同事为代表根据结构主义的观点提出的。斯尼德的著作(1971)一开始的大概思想是,以根据集合论谓词描述理论的框架为出发点,并且运用本章所讨论的公理化类型,然后进展到描述一种理论的概念这一困难得多的问题,理论的概念在本质上是理论的而不是经验的。

关于集合论的保留意见。完全坚持把所有的实体都还原为集合,或者,集合和个体,实际上不是这里所采纳的观点之关键。运用创造的恒等式定义,提出新的、未必是集合的抽象实体,也是自然的,有时更容易。这样的定义之所以称之为是**创造的**,是因为能够通过引入这些定义证明新的命题。在我的关于集合论的书中(1960b),我通过引入个体和运用塔斯基的基数公理,偏离了纯集合论。两个集合 A 和 B 有相同的基数,当且仅当,它们是等价的,即能够进行一一对应。这是一个创造的定义。①

① 关于创造的定义的详细而基本的讨论可参见,Suppes (1957/1999, 第 8 章)。

用符号表示为：

$$K(A) = K(B)\text{，当且仅当，}A \approx B$$

当把这个创造的定义附加到策梅洛-弗伦克尔集合论中时，基数是否是集合，是不可判定的。一个更加初步的例子是，创造序对的定义：

$$(x, y) = (u, v)\text{，当且仅当，} x = u \,\&\, y = v \qquad (1)$$

对我来说，序对的这条进路似乎是最自然和最明显的，尽管库拉托夫斯基(Kuratowski, 1921)和维纳(Wiener, 1914)把序对定义为特定集合是有历史意义的。当新的对象通过像(1)那样的定义创造出来时，我们就能够把它们作为新的个体加到我们的全域中，或者，无论如何作为集合的潜在项，加到新的抽象对象的一个范畴中。看待抽象对象的这种方式将不会使那些希望从一开始就在一个固定的公理集中创造所有这些对象的人感到满意，但我发现，"作为需求的创造"的过程更有吸引力。首先，通过创造恒等式的定义所定义的抽象对象，没有与集合论特性无关的累赘。例如，库拉托夫斯基的定义

$$(x, y) = \{\{x\}, \{x, y\}\}$$

在 x 和 y 所属的集合之间产生了一种不对称。这似乎不是我们的序对概念的一个直观的组成部分。我发现，自然数概念的各种集合论的还原更是如此，因此，前面作为一种创造的定义给出的塔斯基的公理是更令人喜欢的。

尽管如此，这种保留意见在后面将起不到重要作用。我在这里已经概述性地表明，我会接受对待结构概念的一种多元论的态度。现代数学的范畴论(theory of categories)提供了不同类型的另一种论证。在麦克莱恩(MacLane)和伯克霍夫(Birkhoff)(1967)的著作中会找到对这些思想的一种特别容易理解的介绍。 35

2.4 公理化方法的历史视角

在结束本章之前,也许考察一下(即使不很详细)公理化方法的发展史是有趣的。

欧几里得之前。公理化方法至少可追溯到古希腊的数学,这是一个熟悉的历史事实。然而,学者们没有完全理解,在欧几里得时代或先于欧几里得的两个世纪里的发展及其与更早的巴比伦数学和埃及数学相关的故事。公理化方法,正如我们认为它在欧几里得的《几何原本》中是具体的那样,似乎起源于公元前 4 世纪或可能更早一些,主要是由欧多克索斯(Eudoxus)创立的。把泰勒斯看成是几何学之父的传统故事,缺乏现有证据的支持,而且,把在欧几里得几何中表现出的并在柏拉图的《美诺篇》(*Meno*)中已经部分明显的精致的公理化方法归因于泰勒斯,似乎肯定是一种误解。以恰当的视角保持柏拉图的相对适度的作用,也是重要的(Neugebauer 1957,第 151—152 页)。在古希腊数学中,对欧几里得的《几何原本》的历史发展的一个极其详细的讨论,是由诺尔(Knorr)提供的(1975),①但在转向这

---

① 在古希腊数学中对公理化方法的前欧几里得发展的有争议的历史进行的更一般分析,可以在诺尔的著作(1986)中找到。我说"有争议的历史",是因为证据不足。对这种历史的强调应用的更多评论,可以在 6.8 节(第六章的最后一节)中找到。诺尔的评价比诺伊格鲍尔(Neugebauer)的评价更稳妥。诺尔的观点是,在柏拉图学院里,几何学家和哲学家之间有着密切的相互作用。柏拉图的《理想国》(*Republic*)(510C-511C, 527A, 528B-D)中的不同段落,批评了几何学家的严密性。有关基础的这些哲学观点,可能对欧多克索斯等几何学家在他们发展形式方法时产生了影响。但是,问题的另一方面是,像柏拉图那样的哲学家几乎没有对发生在同时代的几何学的许多技术发展产生过影响。根本不存在阻止这些更深层的技术性成果的基础危机。

个工作之前,有必要先讨论一下亚里士多德。

亚里士多德在不同的地方讨论数学的第一原理,但最明确和最详细的讨论是在《后分析篇》(*Posterior Analytics*)中找到的(72a14 - 24, and 76a31 - 77a4)。按照亚里士多德的观点,一门论证性的科学必须从无法证明的原理开始。在《形而上学》(*Metaphysics*)的不同段落中,他特别强调了证明一切的不可能性(997a 5 - 8,1005b11 - 18 和 1006a 5 - 8)。亚里士多德在《后分析篇》中说,在所有这些无法证明的原理中,有些对所有的学科是共同的。这些原理就是公理(αξιώματα)。另一些原理是专门针对某一门特殊学科的。

对亚里士多德而言,公理的标准事例是这种原理:"如果从等量中减去等量,那么,余数相等"和在"两个相互矛盾"的逻辑原理中"一定有一个是正确的"。亚里士多德用其他名字称呼公理,例如,"共同的(东西)"(τα κοινα),或"共同见解"(κοιναι δοξαι)。

至于针对某一门学科的特殊原理,亚里士多德把它们划分为两个部分:首先,断言存在的假设,例如,存在属(genus)或主题(subject matter),在几何学的情形中,是存在量(magnitude)。第二类论题是定义。这样,对于亚里士多德来说,一门学科的第一原理被划分为公理、假设和定义。值得注意的是,尽管亚里士多德对这些思想提供了相当透彻的一般讨论,但是,他没有在任何一个地方举出具有如此阐述的原理的一门学科的详细例子。他确实具有的价值是,以一般而系统的方式给出了对第一原理的最早的广泛讨论。 36

欧几里得。既然关于欧几里得前辈没有更多要讨论的,让我们现在转向对在他的《几何原本》中反映出的公理化方法的稍微详细的讨论。欧几里得在第一卷一开始列出了 23 个定义、5

个公设和 5 个公理或共同概念(Heath 翻译,第二版, 1926)。[①]欧几里得在第一卷中推导 48 个命题时所遵守的精确标准是相当高的。另一方面,正如人们所预期的那样,他的进路与现代的公理化方法概念不完全一致。他似乎没有明确地看到,几何学的公理化发展必须从本身不能用其他概念定义的某些原始思想出发。也许公平地说,在这方面他把形式的或公理的问题与关于几何学的应用问题混淆了。例如,第一卷的定义 1 断言,一个点是没有部分的某种东西。定义 2 断言,一条线是没有宽度的长度,但后来的证明实际上并没有求助于这两个定义中的任何一个定义。与这种情况更接近的是,点和线的概念实际上对于欧几里得来说是原始的,是直觉地加以描述的,而不是通过他作出的明显的和意会的假设在数学上定义的。

在现代的讨论中忽略的另一个区别是,欧几里得在公设和公理之间的区别。普罗克鲁斯(Proclus)在他关于欧几里得的著名评价中关于这种区别是这么说的:

> 它们彼此不同,在这方面,就像定理区别于问题的一样。例如,像在定理中一样,我们根据前提来提出所看到的和所确定的东西,而在问题中,则我们被告知去发现和去做某事,同样,在公理中,人们假定这样的东西是自明的,而且,很容易通过我们的直观概念(untaught notions)来理

---

① 非常有趣的是,欧几里得实际上没有用"公理"(axioms)这个术语,而是只用了"共同的概念"(κόιναδ),尽管含义与亚里士多德的公理或共同的东西之含义是一样的。另一方面,欧几里得用公设(αιτουμενα)替代了亚里士多德的假设术语。尽管欧几里得的影响很大,但是,这个术语在希腊数学中并不是固定的,而是不断地发生变化;例如,在阿基米德的不同著作中发现有希腊术语公设(postulates)、假设(hypotheses)和假定(assumptions)。但在所有这些情形中,从基本意义上看,亚里士多德讨论的假设或欧几里得引入的公设的含义是一样的。

解，而在公设中，我们假定这样的东西是容易发现的和达到的（我们的理解在他们的假定中是没有困难的），而且，我们不需要复杂的工具……两者都具有简单和易掌握的特征，我意指的是公设和公理；但公设命令我们设法发现能呈现出一个简单而容易掌握的特性的某个主题（υλψ），而公理命令我们断言对于学习者来说某种自明的本质属性……

因此，正是几何学家知道，所有的直角都是相等的，如何沿着一条直线截取任意长的线段，而相同的东西也是彼此相等的，这是一个众所周知的见解，而且使一般陈述适合于自己学科的算术家和任何一位科学家都使用了这个概念。①

37

也许正是在几何学专用的公设和对所有研究都是共同的作为假定的公理之间的最后的区别，最好地代表了欧几里得自己接受的观点。尽管很容易挑出欧几里得的《几何原本》中的某些缺陷，也很容易强调他的公理化方法概念与现代的公理化方法概念之间的不同，但基本观点仍然是，在基本概念和方法上，在他的《几何原本》中反映出的公理化方法非常接近于现代的观点。

也许考虑几个跨世纪的历史事例是有益的。由于本书对理论性的科学感兴趣，所以，正如已经强调的那样，我将主要举几个例子，这些例子通常与物理学相关，而不是与纯数学相关，即使从形式的观点来看，两者之间的区别并不明确。这些历史事例的讨论打算突出两个重要的观点。第一个观点是，存在着一个一直延续到牛顿的欧几里得——阿基米德的数学传统，几乎完全独立于从希腊文化到17世纪的哲学家中占统治地位的亚

① 这段引文摘自希思（T. L. Heath）关于欧几里得《几何原本》的介绍和评论（Heath, 1926: pp. 122 - 123）。

里士多德的分析传统。第二点是，物理学中的严格性标准，尽管旨在模仿几何学中的那些标准，实际上根本不如我们一开始大概关于数学物理的第一本著作中（即阿基米德关于静力平衡的著作）定义的标准那么严格。①

阿基米德。我们可以首先考虑阿基米德在他的第一卷"论平面平衡"(*On the Equilibrium of Planes*)的著作中的公设（摘自希思编的阿基米德的著作，1897，第 189—190 页）。

我假设如下：

1. 相同的重体，位于相等的距离，处于平衡态；相同的重体，位于不相等的距离，处于不平衡态，但向着距离较远的重体方向倾斜。

2. 如果当位于特定距离的重体处于平衡态时，在一个重体上增加某物，那么它们就不保持平衡，但向着有附加物的那个重体方向倾斜。

38

3. 同样，如果在一个重体上去掉任何一种东西，它们不保持平衡，向什么都没有去掉的那个重体方向倾斜。

4. 当相等和相似的平面图形相重合时，如果彼此相互应用，那么，它们的重心同样重合。

5. 在不相等但相似的图形中，重心同样在相似的位置。我意指的与相似图形相关的相似定位的点是这样的点：如果从它们画出两条直线，形成相等的角，那么它们成为有对应边的等角。

---

① 但不是关于数学物理的第一本著作。现在广泛认可的一个事实是，那个重大的荣誉归于巴比伦的天文学家们，至少与那波那撒(Nabonassar)统治时期(746—723 BC)和后来的几个世纪。关于巴比伦的天文学家们对历史上第一门数学学科的初期贡献的最近分析，可参见 Swerdlow(1998)的著作。还要强调的是，只有在进行了几千年的观察之后，关于天文现象的详细的数学计算才繁荣起来。

6. 如果特定距离的量保持平衡，那么与这些量相等的(其他)量也在相同的距离保持平衡。

7. 在任何一个它的边界在(一个)相同方向凹进去的图形中，重心一定在这个图形中。

在详细地考察这几个公设的特征之前，关于阿基米德这部著作的历史背景多说几句可能是有用的。这部著作撰写的精确日期确实不知。通常(但没有完全证实)他生活在公元前 287 年到 212 年；他去世的日期似乎是相当可靠的，因为根据几种资料，他是于公元前 212 年在罗马征服了叙拉古(Syracuse)时被杀的。许多希腊数学先于这部著作，但是，它在悠久的数学物理史上占有独一无二的地位，因为它显然是在数学与力学之间建立密切联系的第一部系统的著作。阿基米德阐明的公设现在被看成是可列为**静力学**部分的公设。

关于阿基米德的公理已经有许多方法论的评论。这些评论陷入的许多困难是描述几个世纪以来在试图对物理学的不同分支的基本假定给出一种完美的阐述时产生的典型问题。根据本章迄今阐述的观点，也许阿基米德公理的最惊人的典型特征是它们缺乏完备性。读者立刻会觉得，人们期望的在数学上对这些假定作出的完备说明从没给出。假设某些几何概念本来就直接来自希腊几何——例如，在公设 5 中引入的相似性概念——是公正的，但是，等距概念、平衡概念或特别是重心概念并非如此。无疑，处于平衡态的两个对象关系的基本特性潜在地被假定了，例如，这种关系是可传递的和对称的。最明显的问题是围绕一个平面图形的重心概念提出的。这个概念对于公设 4、5 和 7 的表述来说是必需的，但另一方面，这些公设本身很显然没有为这个概念提供一种完备的描述。在此我的意思是说，如果我们除了公设 4、5 和 7 中的陈述之外对重心概念一无所知，那么

我们就不可能得到阿基米德感兴趣的而且他确实已经得出的定理。正如狄克斯特休斯(Dijksterhius)在他关于阿基米德的书(1957)中所指出的那样,有可能认为,阿基米德已经从更基本的讨论中接受了重心概念,因此,在他的著作中,重心概念实际上与几何的相似性概念占有相同的地位。从表面上判断,这种论证似乎比托普利茨(Toplitz)和斯泰因(Stein)的下列论证(Stein, 1930)更可靠:他们提议,一旦这些公设被明显的和自然的假定所扩大,那么,它们将被看成是明确地定义了重心概念。

39

在本章第二节阐述的意义上,无疑不可能以任何简单确切的方式给出阿基米德理论的一种标准形式化。给出在公设 1、2、3 和 6 所包含的理论部分的一种标准形式化是有可能的。(关于这是联合测量理论的第一个公理化事例的一种论证,可参见 Suppes 1980。[①])在根据阿基米德公设的集合论术语提供的一种重构中,存在着严重的问题,更何况标准形式化的问题,对于像静力学那样的学科来说,标准形式化的问题是我们几乎料想不到的。在这样一种集合论的表述中,我们能够毫无困难地使用像相似性那样的几何概念。如果我们接受事先提出的重心概念的定义,那么,例如,从这些较早的发展来看,公设 4 似乎只是一个定理,没有必要进行个别陈述。换句话说,根据对重心概念的这种处理方式,阿基米德理论的原始概念没有一个出现在公设 4 中,因此,它显然是一个可排除的公设。同样的评论适用于公设 5 和 7。阿基米德只解决了半个问题;他的公设没有对重心概念提供一个完备的描述,但另一方面,也不能把这些公设说成是依赖于对重心概念的一个完全独立的描述。

---

① 3.4 节的定义 9 给出了联合测量的基本公理。

这些困难是从一开始就困扰着物理学和其他经验科学的基础研究的那类困难。几何学的情形是靠不住的。几何学能够建立在相当少的概念基础之上，而且，即使欧几里得的《几何原本》不代表一种完美的公理化发展，提供一种完全严密的发展也是今天最直接的问题。在阿基米德关于静力学的著作中，正如在对经验世界进行物理学处理的所有情形中那样，有必要引入比几何学概念更多的概念。不顾公设 1、2、3 和 6 的基本性质，引入的许多概念，从形式的观点来看，必须被看成是原始的。考虑到公理的简单性，这个列表是令人生畏的：具有相同重量的两个物体的概念；比较距离（从一个支点出发）的一种关系；平衡的一种二元关系，或起码的一种二元弱排序关系，我们可以根据这种关系来定义平衡和倾斜关系；最后，对组合物体的一种二元运算是阐述公设 2 和 3 所需要的。对阿基米德来说，正如对他的大多数数学物理方面的继承者一样，对其中许多具有十分明确的直观含义的概念作出明显的分析，无疑是没有必要的。但是，关于直观含义的这种评论过分简单化了。如果彻底地追踪这条思路，非凡的希腊人关注的严密性和数学的演绎结构将成为一个无法说明的谜。也许设想这部著作是以相当高的水平写成的，并且假定了关于这个主题的许多前期分析是合理的。确定在多大程度上这是真的需要比在当前语境中的可能情况作出更多的文本分析。 40

阿基米德的著作决不是一种孤立的情形。其他优秀的古代著作经常不陈述公理，而是一开始就给出一系列定义。两个例子是，阿波罗尼斯（Apollonius）的著作《圆锥曲线》（*Conics*），这本书以其数学的严密而著名，以及阿基米德自己的小册子《论螺线》（*On Spirals*）。在同样的严密程度上，一个更加物理学的例子是狄奥克勒斯（Diocles）的《论燃烧的反射镜》（*On Burning*

Mirrors)。似乎正确的情形是,这些著作的特色都是毫无疑问地接受了公理化框架和欧几里得的《几何原本》所规定的演绎的严密程度。对《几何原本》的这种广泛接受,在惊人的程度上,对应于20世纪类似于对有选择公理的策梅洛-弗伦克尔的集合论(ZFC)的接受,或者,对作为几乎所有的现代纯数学的特有框架的ZFC的某种次要变种的接受。一个初步的阐述可以在苏佩斯的著作(1960b)中找到。

在结束关于阿基米德著作的讨论时,值得注意的是,马赫(Mach)在他著名的力学著作中,似乎严重地混淆了阿基米德的证明。马赫讨论的焦点是阿基米德书中的公设6,这个公设断言,可通约的量处于平衡态时,距离与它们的重量相互成反比,这恰好是阿基米德对著名的杠杆原理的阐述。马赫(1942,第20页)尤其关心这样的事实:"(这个命题的)完全演绎包含了通过假定(如果不是明确的话)来证明的命题。"马赫所混淆的一个核心观点似乎是,他完全误解了把数学应用于物理学的本性。马赫似乎完全不懂,如何用数学从一般假定中推论出特殊命题,以及这些一般假定与特殊命题之间是什么关系。他似乎错误地以为,任何一个像刚才引用的那个命题一样的命题一定是以某种方式从经验中直接建立起来的。他关于这些问题的观点在下面一段中清楚地表达出来:

> 重量和杠杆臂成反比,是从等距等重的纯粹平衡的假定中得出的!这如何可能呢?如果我们不能从哲学上先验地研究平衡依赖于重量和距离的简单事实,而是被迫努力从经验中获取那种结果,那么,在多么低的程度上,我们能通过思辨的方法发现这种依赖关系的形式,即比例关系!
>
> (Mach, 1942: p.19)

马赫关于数学与物理学关系的错误观点，对整个 20 世纪的物理哲学的讨论，产生了非常不幸的影响。

为了表明阿基米德所缺乏的明显的公理的严密性是古代科学中其他理论著作的特征，我简要地讨论另外两个古代的例子。

欧几里得的《光学》。强调下面这点是重要的，即欧几里得的《光学》(*Optics*)实际上是一个视觉理论，并不是一本关于物理光学的著作。从单眼视觉(monocular vision)透视的观点来看，许多命题都与视觉相关。确实，欧几里得的《光学》能够被描述为是在欧几里得几何范围内关于透视图的一本专著。欧几里得著作的基调能够从开始部分的引文中看得出来，这段引文由七个"定义"组成：

41

1. 允许假定，直接从眼睛划出的各种不同的线通过很大的空间范围；

2. 在我们的视线范围内包含的空间形式是一个圆锥体，这个圆锥体的顶点在眼睛上，底部在我们视线的界线；

3. 我们的视力能及的那些东西是能被看见的，而我们的视力不及的那些东西是不能被看见的；

4. 在较大角度的范围内看那些东西好像较大，在较小角度的范围内看那些东西好像较小，在相等角度范围内看那些东西好像大小相等；

5. 在较高视线范围内看那些东西好像较高，而在较低视线范围内看那些东西好像较低；

6. 同样，在右眼的视线范围内看那些东西好像在右边，而在左眼视线范围内看那些东西好像在左边；

7. 但是，从几个不同角度看到的那些东西好像更清楚。

[这段引文引译自伯顿(Burton)1945 年的英文译本]欧几里得《光学》的发展在特征上是数学的,但它的公理化与《几何原本》的公理化不一样。例如,欧几里得后来证明了两个命题,“知道在阳光灿烂时一个指定的海拔有多高”,以及“知道在太阳落下时一个指定的海拔有多高”。正如将要预期的那样,重要的问题是根据太阳落下光线会终止作为光线的证明,除了以这种非正式的、常识的、物理的方式处理这两个命题之外,根本没有认真地提出太阳或光照的概念。视觉空间当然是被欧几里得在特征上作为欧几里得的空间来处理的。

可能有人提出反对说,在欧几里得的《几何原本》中也存在着类似的形式缺点。但他们没有经过更多的思考就承认,在《光学》的这些定义中许多物理学术语的引入和在《几何原本》中发现的很有限的语言用法之间有很大的不同。似乎对我来说,欧几里得《光学》中对基本假定的阐述,正是受了所谓的(在我们自己时代的)物理学公理的精神的很大鼓舞。没有在任何类型的数学严密性上做出努力,而是努力从直觉上传达这些根本的假定。

托勒密的《天文学大成》。我引证的第三个也是最重要的例子是托勒密的《天文学大成》(*Almagest*)。之所以说它是重要的,是因为它是古代最重要的科学著作,也因为它并不包含任何伪装的公理化处理。托勒密确实使用了数学论证,而且的确是很熟练的数学证明,但他是以一种应用的方式运用数学的。他没有引入关于星体运动的明确公理,但把关于星体运动的研究还原为几何命题,当然包括了球面三角学的重要例子。

下面《天文学大成》接近开头的一段很好地举例说明了托勒密据以提出假定的精神:

> 这种一般的初步讨论涵盖下列主题:天体在形状上是

> 球形的，并像一个球体那样运动；当把地球作为一个整体来看待时，它在形状上也显然是球形的；它位于天体的中间位置，非常像是天体的中心；在大小和距离上，它相当于一个点与固定恒星的球面之比；而且，它不会从一个地方运动到另一个地方。
>
> (H10, Toomer translation, 1984: p.38)

42

接下来对这些问题逐一进行了较长的和更详细的讨论，例如，天体按照圆形轨道运动。我的论点是，这种讨论和讨论的框架完全接受了我们今天所认为的非公理化的理论科学。根本没有以公理化的方式整理这些思想的任何迹象。

当托勒密认真考虑细节时，他作出了下列论述：

> 我们现在开始进行单独证明，我们认为，首先应该沿着它们画出的一个大圆圈来确定上面提到的（黄道和赤道）两极之间弧线的大小。但我们看到，首先有必要说明确定弦的方法：我们将在几何学的意义上彻底地图示所有的论题。
>
> (H31, Toomer translation, 1984)

因此，关于圆周上内切弦的大小的详细计算，首先是强调计算，而且它的风格还有结果使现代物理学家感到高兴。这种冗长的和重要的计算分析包括了弦的数值表。

第四卷中对月球运动的处理，在许多方面甚至更引人注目地举例说明了我所拥护的论点。在这里，托勒密关注适合于研究月球运动的那种观察，而且特别关注各种不同的观察如何得到矫正和如何提出一致性理论的一种方法论。

后来的书中提出的各种假设，例如，第五卷中月球的加倍异常假设，是接受了现代天文学或物理学，不是公理化的数学的精

神。此外,托勒密在《天文学大成》一书中对几何定理和证明的自由而有效的运用,似乎格外地类似于在数学物理的某个领域的现代著作中对微积分和微分方程理论的运用。①

在对运用公理化方法以及从像光学和天文学之类的古代理论科学来看他们缺乏明显形式的这种分析中,我只部分地进入了关于公理、公设和假设地位的哲学分析的讨论中。从前面引证过的柏拉图到普罗克鲁斯,有大量的古代文献论述这些问题。也许,尚存的最可靠和最认真的讨论可以在亚里士多德的《后分析篇》中找到。亚里士多德以十分明确和有说服力的方式说明了,几何学的证明如何能够被适当地用在力学和光学中(75b14ff)。但正像亚里士多德实际上根本没有从光学、力学或天文学中举例一样,似乎对我来说,他作出的令人感兴趣的那些区分,同样不能帮助我们更好地理解欧几里得关于他的光学定义的观点,或者阿基米德关于上面引用的重心的观点。

43

约旦努。阿基米德运用的这种公理化进路对中世纪的物理学讨论产生了极强的影响。对这个时期更全面的力学史感兴趣的读者可参考克拉格特(Clagett)的力学卷(1959)。我在这里只考虑一个重要的例子。与认为中世纪缺乏认真的科学与数学发展的某些流行观点相反,20 世纪的学者们,从迪昂(Duhem)

---

① 有所保留地说,欧几里得关于天体天文学的公元前 4 世纪的著作《观测天文学》(*Phaenomena*)也能够这么说(Berggren and Thomas 翻译, 1996)。顺便一提的是,从所作出的天文学的假定来看,命题 1 作为一个严密的数学证明,在古代世界中是著名的,引文是"地球处于宇宙的中间,并占有着宇宙的中心位置"(第 52 页)。几百年后,大约到了公元 130 年,盖伦(Galen)在他的《论希波克拉底和柏拉图学说》的著作中这么说这个结果:

> 欧几里得在《观测天文学》的定理 1 中用很少的几句话表明,地球作为一个点或一个中心,位于宇宙的中间,并且学生们相信这种证明,好像它是二加二等于四。(Berggren and Thomas,第 8 页)。

开始,自 1950 年以来,以马歇尔·克拉格特(Marshall Clagett)及其合作者与同事为先锋,已经证明并非如此。现在这个小组编辑和翻译的中世纪的许多著作足以明确地表明了他们写了多少部著作。此外,在许多情形中,所运用的方法是接受了欧几里得和阿基米德的公理化方法。在这些文献中,有一大批著作论述重体的理论,在穆迪(Moody)和克拉格特的著作中,公布了其中的一个重要实例。或许,论重体的这些著作中最重要的一本是由 13 世纪的巴黎数学家约旦努(Jordanus de Nemore)撰写的《重体测量篇》(*Liber de ratione ponderis*)。

按照长期存在的传统,这些公设,特别是约旦努的定理,充分运用了欧几里得几何的经典概念和定理——没有明确地这样说。下面是对约旦努的七个假设或公设的一种英文翻译(Moody and Clagett, 1952: p. 175):

公　设

1. 每一个重体的运动都是向着(世界的)中心的,而且它的力是趋向于下落运动和阻止反方向运动的力。

2. 较重的那个物体下落较快。

3. 较重的物体在下落时,在某种程度上,更直接地向着(世界的)中心运动。

4. 处于适当位置的较重的物体,在那个位置时,它下落的轨道几乎是垂直的。

5. 一种更加倾斜的下落是在相同距离内不垂直的下落。

6. 如果一个重体的上升是由另一个重体的下落引起的,那么,在适当的位置,这个重体比另一个重体较轻。

7. 相等的位置是指与垂线的夹角相等的位置,或使得这些角成为直角,或使得杠杆平行于地平面。

我下面引用的只是第一部分的第二个定理，它是特别令人感兴趣的，因为它对悬挂相等重体的等臂的稳定的水平平衡进行了明确的阐述（和证明）。

定理 2. 当一个等臂平衡的杠杆位于水平位置时，那么，如果把相等的重体悬挂在它的两端，它将保持水平位置，如果把它从水平位置移开，它将会恢复到水平位置，但是，如果悬挂不相等的重体，平衡将落向较重的那一边，直到它达到垂直位置。

44

牛顿。我下面考虑在现代科学史上可能是最重要工作的初始定义与公理。我想强调的主要历史观点是，牛顿是在欧几里得的《几何原本》的希腊传统中撰写了他的《自然哲学的数学原理》。他运用的抨击的数学方法，乃至更多的阐述风格，都与那个传统非常相符。在考虑阿基米德或牛顿还有欧几里得时，重要的不是被公理和定理的言语阐述所阻碍。这再一次体现了一个很好地建立起来的希腊数学传统。即使这些作者的数学著作在很大程度上是用口头术语来阐述的，但是，在他们的阐述和在亚里士多德的《物理学》中以及与来自亚里士多德而不是阿基米德和欧几里得的中世纪传统中发现的阐述之间，有着明显的概念上的不同。在亚里士多德的传统中撰写的著作在特征上真的是非数学的。他们的基本假设没有任何普通数学的阐述，也几乎没有对结论的数学证明。从方法论的观点来看，笛卡儿于 1644 年第一次出版的著名的《哲学原理》（*Principles of Philosophy*）属于亚里士多德的传统，而牛顿于 1687 年首次出版的《自然哲学的数学原理》则属于欧几里得和阿基米德的传统。

人们通常断言，现代物理学的数学方法起源于 17 世纪，而且是在中世纪的亚里士多德传统的曲折演化之后产生的。事实

肯定不支持这样一种观点。受欧几里得和阿基米德的影响,存在着一个连续的数学传统,范围从中世纪一直到 17 世纪。当然,这条数学进路的连续性也强烈地影响了伽利略。也许最明显地表明这种连续性的物理学分支是天文学。① 观察天文学和数学天文学的历史在本质上有一条连续的发展线索,可追溯到前希腊巴比伦的天文学家,而且,在任何一个时期都很少受亚里士多德的分析模式的影响,尽管托勒密接受了亚里士多德的一般宇宙学。

牛顿在他的《自然哲学的数学原理》中一开始提供了下列八个定义和三个定理或运动定律[我引用了卡乔里(Cajori)于 1946 年的翻译]。

定义Ⅰ 物质的量是由它的密度与体积结合起来度量的。

定义Ⅱ 运动的量是由物质的量(quantity of matter)和速度(velocity)结合起来度量的。

定义Ⅲ 所谓 vis insita,或物质固有的力是一种抵抗力,每一个物体由于这种抵抗力,尽可能地保持它当前的运动状态,无论它是静止状态,还是在一条直线上的匀速运动状态。

定义Ⅳ 外加的力是一种为了改变一个物体的静止或匀速直线运动状态,而外加于其上的一种作用力。 45

定义Ⅴ 向心力是推动或拖曳物体,或以任何方式趋向一个中心点的力。

定义Ⅵ 向心力的绝对量是由正比于引起它从中心传向周围空间的原因的效力来度量的。

---

① 一个重要的例外是,惠更斯(Huygens)的《摆钟》(*Pendulum Clock*)(1673/1986)。很容易论证,在传统几何风格中写成的数学物理的最近两本重要的原始著作是惠更斯的这本书和牛顿的《自然哲学的数学原理》。

定义Ⅶ 向心力的加速的量是由正比于在特定时间内产生的速度来度量的。

定义Ⅷ 向心力的运动的量是由正比于在特定时间内产生的运动来度量的。

定律Ⅰ 每一个物体都继续保持其静止或匀速直线运动状态,除非有外力作用于它,迫使它改变这种状态。

定律Ⅱ 运动的改变与所加的动力成正比,并且发生在所加外力的那个直线方向上。

定律Ⅲ 每一种作用总是有一种相等的反作用:或者说,两个物体彼此间的相互作用总是相等的,并且指向对方。

我们应该马上注意到,这些定义与定律每一个都伴随着附加的评论和说明。从方法论的观点来看,重要的是,牛顿《自然哲学的数学原理》的最初的系统性与欧几里得的《几何原本》是多么接近。正如在欧几里得的情形中那样,牛顿也是从一系列定义开始的,试图根据非常熟悉的概念来定义他的关键概念。像已经评论的那样,从形式公理化的观点来看,这样的定义是不适当的。另一方面,它们肯定确实起到了提供预期的经验意义的作用。然而,这些定义所起到的作用,比简单地提供这样一种经验解释的作用更系统,因为它们以一种系统化的形式方式把这个理论的形式陈述中主要提出的各种不同概念联系起来。在欧几里得或牛顿的著作中,理论结构的这两个方面没有完全分离开来。①

---

① 《几何原本》和《自然哲学的数学原理》确实完全不同,牛顿在《自然哲学的数学原理》中包括了详尽的天文学数据(第三卷)。乃至某些与流体力学而不是天文学相关的关于流体方面的数据(第二卷)。在这方面,《自然哲学的数学原理》比《几何原本》更接近于托勒密的《天文学大成》。另一方面,第一卷的严密的公理化和数学推导在精神上与《天文学大成》相比更接近《几何原本》。

现代几何学。现代的公理化方法的历史根源是对 19 世纪几何学基础的极其详尽的审查。无疑,支持这种努力的最重要的驱动力,是由鲍耶(Bolyai)、洛巴切夫斯基(Lobachevski)和高斯在 19 世纪初发现和提出的非欧几何。欧几里得的《几何原本》中严重的逻辑缺陷也受到持久的关注。19 世纪初期,例如,庞斯列(Poncelet, 1822)提出了他的相当不能令人满意地阐述的连续性原理,他能够运用这个原理证明,如果在一条直线上有一个圆里面的一个点和外面的一个点,那么它一定有与这个圆相交的两个点——从欧几里得的公设不可能证明的一个结果。后来表明,这个结果可根据狄德金(Dedekind)的完备性公设(1872/1963)得出,这个公设有一个令人满意的阐述: 46

> 如果一条直线的所有的点都落在两个类当中,使得第一类中的每一个点都位于第二类中的每一个点的左边,那么就存在着惟一的一个点,这个点把所有的点区分为两类,把这条直线划分为两个部分。
>
> (Dedekind, 1872/1963: p. 11)

帕施(Pasch)的公理(1882)是一个更直接的几何公理,因而也是从直观上更引人注目的公理,它填补了欧几里得公理的空白。这个公理断言,如果一条线与一个三角形的一条边相交,那么它也一定与第二条边相交。

首先,正是帕施在他 1882 年的有影响的著作中示范了现代形式的几何概念。[对这些发展的一个极其详细的讨论可在内格尔(Nagel,1939a)的文章中找到]。下面是帕施关于几何概念的意义作出的论述(引自内格尔文章的翻译):

> 确实,如果几何学真是演绎的,那么这种演绎必须处处与几何概念的含义无关,正像它必须与图形无关那样;只有

在所用命题与定义中所详细说明的关系，才能合理地加以考虑。在演绎的过程中，思考这些术语的含义，是有用的和合法的，但决不是必要的；事实上，如果有必要这样做，那么，作出这种证明显然是不充分的。然而，如果一个定理是严格地从命题集合——**基本的**集合——中推导出来的，那么这种演绎具有超越了最初目标的价值。因为如果根据用特定的其他术语来替代这个基本命题集合中的几何学的术语可以获得真命题，那么也可以在这个定理中作出相应的替代；这样，我们由于改变了基本命题而获得新的定理，不必重复这种证明。

(pp. 237 - 238)

帕施的观点在所有的基本方面都与上一节略述的公理化理论的集合论进路相一致。

意大利几何学家皮埃里(Pieri)和韦罗内塞(Veronese)以及佩亚诺(Peano)广泛地评论与推广了帕施的形式公理化进路。但显然，在这条发展线索中，最有影响的工作是希尔伯特(Hilbert)首版于1897年的《几何基础》(*Grundlagen der Geometrie*)。希尔伯特不仅是当时最杰出的欧洲数学家之一，而且他以十分清晰但基本的方式，最低限度地运用了数学符号撰写了他的短篇论著。在几何学的历史上——至少在20世纪，这可能是欧几里得之后得到最广泛引用的著作。但是，重要的是牢记，实际上是帕施，而不是希尔伯特，第一个对几何学的现代公理化进路提供了明确的阐述，这种阐述在现代数学的发展中已经产生了同样的影响。

当公理化方法在概念上和逻辑上已经很成熟时，尤其是在帕施和希尔伯特的著作中，对19世纪的几何发展史的一种简单浏览，提出了关于公理化方法发展的一个自然问题。这种公理

化关注的基础为什么不是投影几何，而是非欧几何呢？从现代的观点来看，这似乎是奇怪的。例如，在博尔舒克（Borsuk）和 47 斯米卢（Szmielew）（1960）的精确的公理化中，在欧氏几何需要的很长的公理单上只改变一个公理，我们就获得了双曲几何，即最重要的非欧几何。相反，投影几何的任何标准的现代公理化看起来感觉完全不同于欧几里得的公理化。但同样，对这个问题的这种回答是简单的。与庞斯列、夏斯莱（Chasles）、斯坦纳（Steiner）和冯·施陶特（von Staudt）的工作相联系的投影几何发展的初期乃至繁荣时期，仍然停留在欧几里得的分析框架之内。他们典型地研究所谓在分析的意义上描绘出的图形的图示特性。例如，庞斯列研究不变的特性，这些特性不是使距离或角度的大小保持不变，而是在其坐标的投影变换下所保持的不变。[①]

希尔伯特和弗雷格。无论如何，如果没有它们的哲学评论家，特别是弗雷格，就不会有这些发展。尽管弗雷格与帕施有信件往来，但是，弗雷格的信件有一部分丢失了。相比之下，有幸保存下来的弗雷格与希尔伯特之间的信件已经得到了广泛的讨论。下面是弗雷格在 1899 年 12 月 27 日写给希尔伯特的一封信中所表达的批评希尔伯特的《几何基础》的核心部分：

> 对第 1 部分和第 3 部分的说明显然是完全不同的类型，因为“点”、“线”、“之间”这些词的含义并不是给定的，而是假定预先已经知道。至少看起来如此。但是，你所称的一个点是什么，也还是不明确的。人们首先在欧几里得几何的意义上考虑点，这是公理表达了我们直觉的基本事实

① 这种评论附加在庞斯列与朱尔·维耶曼（Jules Vuillemin）的讨论之后。

> 这个命题所强化的一种思想。但之后,你把一个数对看成是一个点。我对下列命题持有怀疑态度:对关系的一种精确而完备的描述是由几何公理提供的(第1部分),以及“之间”这个概念是通过公理定义的(第3部分)。这里提出的公理承担了属于定义的责任。对我来说,这似乎以一种不确定的方式消除了在定义和公理之间划出的界线,而且除了“公理”这个词的旧的含义(这个含义出现在下列命题中:公理表达了直观的基本事实)之外,还出现了我不再能够完全把握的另外一个含义。
>
> (Frege, 1899/1980: pp. 35 - 36)

弗雷格在他的长信的稍后评论说,“因此,公理和定理决不能努力确定其中出现的一个符号或词的含义,而是这种含义必须是已经确定的。”

希尔伯特在两天后,即1899年12月29日,写给弗雷格的回信中以坚定的不同意的口气作出了回应。这封信的核心的一段是:

> 你写道:“我把公理称为命题……从这些公理的真理得出的结论是,它们不是彼此矛盾的。”我发现,在你的信中,读到这一句时是非常有趣的,因为只要我一直在思考、写作和讲授这些东西,我就一直说恰好相反:如果任意给定的公理与它们所有的推论都不相互矛盾,那么它们就是真的,而且公理定义的东西是存在的。对我来说,这就是真理和存在的标准。
>
> (Hilbert, 1899/1980: pp. 39 - 40)

关注这种差异的一种方式是,弗雷格反对希尔伯特对欧几里得平面的分析模型——用数对表示点——来证明公理的一致

48

性。弗雷格无法承认这是一个可接受的公理模型。他内心里的观点与我所采纳的集合论的观点相差甚远。几何的公理应该只适用于"真正"的几何学的对象。

后来在 1900 年 1 月 6 日的信中,弗雷格勉强承认,希尔伯特的观点能够被看成是一种"高瞻远瞩的观点,从这种观点来看,欧几里得几何好像作为更具综合性的理论结构的一种特殊情况……"(Frege, 1900/1980: p. 43)。弗雷格在对希尔伯特关于他的各种公理的独立性证明进行评论时,作出了这种陈述。刚才引用的弗雷格的陈述,似乎支持了集合论的进路,但是弗雷格还是怀疑这条进路的正确性。

同样必须指出的是,希尔伯特对自己观点的表述并不像所渴望的那样明确,但是他的观点和直觉成为从事几何基础研究的数学家们的观点与直觉。下一章,在扩展了这种历史视野的地方,考虑在这种通信中和其他需要表征定理时所缺少的明确讨论。

这次通信之后,弗雷格发表了关于几何学基础的几篇文章(1903a, b; 1906a, b, c),在这些文章中,他更详细地阐明了他写给希尔伯特信中略述的思想。根据通常的标准,这些文章事实上是冗长的、很不详细的和重复的。尽管如此,弗雷格确实最后比希尔伯特更清楚地阐述,希尔伯特对欧几里得几何的公理化处理如何能完全被看成是二流(second-level)的形式理论。然而,就数学中的一般用法而言,正是帕施和希尔伯特的公理化方法的进路,而不是弗雷格的进路,最终获得了胜利。

物理学。在阿基米德或牛顿的情形中,数学和数学物理之间并没有明确地分离开来;这将被看成是现代的区分。从公理化方法的观点来看,在历史上令人感兴趣的是,17 世纪的数学物理是作为数学论述的一部分,并主要以相同的几何方式来撰

写的。自牛顿的《自然哲学的数学原理》出版之后的近三个世纪内,数学和数学物理中所用的演绎方法之间的分歧已经变得相当明显。数学越来越向着明确定义的精确标准的方向发展;上一节略述的这条集合论进路现在是已出版的大量数学研究的典型进路。很自然,我们会发现,在集合论定义的言语阐述和风格方面有细微的差别;但在几乎所有的情况下,集合论描述的重构是一件常规之事。

这完全不是物理学中的情形,即使这里数学物理与理论物理之间是有区别的。数学物理越来越成为数学家研究的一门学科,而且明显地关心假定的明晰性和严密性。另一方面,理论物理的研究方式完全达不到令人满意的数学标准。在本世纪,历史上的相对论或量子力学的重要文章没有一篇是以明确地描述的公理化方式写成的。[说到相对论,我意指爱因斯坦和洛伦兹(Lorentz)的早期论文,不是明科夫斯基(Minkowski)、维布伦(Veblen)、罗布(Robb)等人后来的数学工作。①]甚至冯·诺伊曼(von Neumann)关于量子力学的著作,也没有给出量子力学的公理化发展,只不过是希尔伯特空间的公理化发展。

49

物理学家为了取代数学的严密性奇妙地运用的物理学直觉,一直没有得到适当的研究。这方面的一个好的事例是,爱因斯坦对哥德尔建构的爱因斯坦的引力方程的周期宇宙解——它原则上允许"在一个足够宽的曲线上乘坐宇宙飞船进行一次往返旅行……旅行到过去、现在和未来的任何一个区域后再返回……"——的答复(Gödel, 1949: p. 560)。爱因斯坦的答复如下:谈到哥德尔的宇宙解,"令人感兴趣的是权衡一下,由于物理学的理由,这些是否是不可能排除的"(Einstein, 1949:

① 关于相对论的进一步相关的细节,参见6.3节。

p. 688)。

物理学集中于解决问题。当数学的严密性成为障碍时，它就像其他概念或技巧一样得不到重视。这种实用的态度在我早期的文章中曾进行过更广泛的讨论(Suppes, 1998)。也参见5.8节物理学家关于概率本性的实用态度的分析。

预言公理化方法在经验科学中的总的未来是很困难的。经济学家广泛地使用了这些公理化方法。无论如何，公理化方法现在都被广泛地用于特殊学科的基础研究中，也用来追踪方法论的某些一般问题，特别是关于概率、统计和归纳的那些问题。这些方法的使用，允许我们把成为数学科学公认部分的清晰性和严密性标准带到科学哲学当中。一种保守的预言是，即使在这个世纪它们被非形式的哲学分析或科学分析的语境包围，也还能把它们继续应用到整个这一世纪的基础性工作中。形式的和非形式的这样一种混合在后面几章都会发现。在下列意义上这是令人渴望的和必要的，即科学哲学中的许多有意义的思想，并不是以一种使它们适合根据系统化的公理得到表述的方式得以表述的。这些事例绝大多数来自科学的所有部分中的实验和统计实践。但同意这一点——我认为这样做是重要的——通过尽可能完全地考察满足不同理论的明确的理论公理之结构的表征和不变性，在澄清理论本性时，仍然还有许多有用的工作需要去做。

3.

同构表征理论

51 科学哲学中的一个核心论题是分析科学理论的结构。我自己的大多数工作都与这个论题相关,不过是以一种特定的方式相关的。我多年来提倡的基本进路是,用理论模型分析理论的结构。一般情况下,洞察复杂理论结构的最好方式是寻找其模型的表征定理,因为一个复杂理论的句法结构通常很少提供关于理论本性的理解。我在这里以一般方式提出这种思想,并阐述我早期论文中的观点。我在第一节首先对表征的本性进行了非形式的评论。第二节专门讨论理论模型同构的核心概念。第三节聚焦于某些简单而有意义的表征定理。第四节考虑数值表征的某些基本测量事例。第五节关注一个重要的形式结果,即通过一个无限寄存器机或通用图灵机的程序来表征任何部分递归函数的结果。第六节在某种程度上带有历史性,转向关注心理表征的哲学观点,以及它们在多大程度上需要某种同构概念。

3.1 表征的种类

对某物的一种*表征*是一种意象(image)① 、模型或对那个东西的复制。谈论表征在日常话语中是常见的和频繁的。② 一些

典型的例子如下：

睡觉是死亡的一种特定的意象和表征。

这出戏剧是我曾经很熟悉的世界的一种表征。

它正好是地上乐园的一种表征。 52

在这幅画中，对阿基里斯(Achilles)的表征是不可思议的。

这是对奥古斯都(Augustus)建立的凯旋门的一种表征。

不可能对核力进行一种直观而看得见的表征。

在某些情况下，我们能够把一种表征看成是改进我们对所描绘对象的理解。我们许多人在看了大楼的建筑图之后，肯定更好地理解了大楼的比例——特别是楼内布局。

表征的形式理论或数学理论具有的一个主要目标是丰富我们的理解，尽管表征的其他目标也几乎是同样重要的——例如，运用测量程序的数值表征会使计算更有效。在这里所阐述的形式的意义上，表征也与还原密切地联系在一起。几乎所有的人都接受的一个令人敬佩的目标是，把未知的还原为已知的。当还原的主张在特性上是意识形态的而不是科学的时，争论就产

---

① image这个词在文中有不同的翻译，早期的哲学著作中通常译为“图像”，在科学哲学中有的译为“形象”或“映像”，在近来的心理学的著作中译为“意象”。为了与现在的文献统一，本文在偏哲学的经典文献方面译为“图像”，在偏心理学的文献方面译为“意象”。——译者

② 这里不分析表征的其他含义，即使人们会发现其中的许多含义很类似，就像在工会代表昨天着手管理中的含义一样。在刘易斯(G. H. Lewes)对夏洛特·勃朗特(Charlotte Brontë)的小说《简·爱》的评论(1847)中，令人满意地描写了客观表征和主观表征之间的一种熟悉的对比，“在她对乡村的房子和上流社会的描绘中，有一个人很容易且准确地懂得她在描述什么……这种客观表征的才能也与一种主观表征的奇特的能力结合在一起。我们不仅意指激情的能力——艺术家的心理直觉，而且还意指把外表与代表物质现象的心理解释的内在效果联系起来的能力。这一点在许多精致的描述中体现出来；但是，我们选择了把孩子关在旧的卧室里惩罚的描述，因为这同时显示了我们所谈的能力和前面提到的代表事物的物质方面的能力。”

生了。人们通常不重视,如何把科学的一个部分——甚至是前后相邻的部分——相关地和在技术上现实地还原为另一部分的问题。

关于把一类现象或一组想法还原——因而表征——为另一类现象或另一组想法的哲学主张,与哲学本身一样古老。下面是伊壁鸠鲁主张把一切都还原为简单物体,即原子和空间的论述:

> 此外,宇宙是物体和空间:因为在所有人的经验中,感觉本身证明,物体是存在的,而且,正如我所说的那样,依照感官证据,我们一定有必要通过推理作出关于不可感知之物的判断。如果没有我们所称的空虚、空间和无形的存在,那么物体就没有地方存在,也没有任何东西,像我们所看到的它们的运动那样,借此运动。此外,这两个虚无都能用概念来思考,或者根据与可想到的东西的类比来考虑,比如,能够作为整体的存在来把握,而不能作为这些存在的特性或偶然性来谈论。而且,在物体中,有些是复合体,另一些形成了这些复合体。并且,后者是不可分割的和不可改变的。
>
> (Epicurus, Oates edition: p. 4)

引自伊壁鸠鲁的这段话大约写于公元前 300 年,卢克莱修大约在 250 年之后所写的长诗《物性论》(*De Rerum Natura*)中在几个地方差不多转述了这段话(Oates, 1990: pp. 69 - 219)。把所有的现象都还原为虚空中的原子运动是古代原子主义的核心论点,[①]

① 为了与 20 世纪发展起来的物理学中的"原子论"(the theory of atom)区别开来,这里把"atomism"翻译为"原子主义",没有翻译为"原子论"。——译者

这些思辨思想的发展，对更后来产生的科学发展具有重要的意义。

笛卡儿的几何向代数的还原，是非常接近于这里倡导的形式精神所要求的一种还原，也是思想史上重要的一种还原。他在其《几何学》(*La Geometrie*)(1637/1954, p. 2)一书的开头部分关于这个问题是这么说的：

> 几何学中的任何一个问题都能被轻易地还原为这样的术语：知道一条特定直线的长度足以划出这条线。正如算术只是由四五种运算，即加、减、乘、除和开方(可以被认为是一种除法)所组成那样，在几何学中，为了找到所要求的线，只需要加减其他线段；要不然，画一条我称之为单位线段的线，以便使它尽可能密切地与数字联系起来，而且这条线是任意地划出的，然后给定另外两条线，寻找第四条线，这条线与给定的一条线之比，等于给定的另一条线与单位线段之比……

53

这两种还原的论点，在当初它们的概念提出的意义上，它们之间的区别几乎不可能再大了。关于它们的还原论题，古代原子主义者能够在令人满意的科学意义上建立他们的还原论点，但这个论点在实践中是无法证明的。笛卡儿详尽的数学处理成为现代数学初期的一个重要的概念突破。另一方面，笛卡儿在他的《哲学原理》(1644)一书中把物质只还原为外延的企图，在其方式上，恰好与伊壁鸠鲁或卢克莱修的企图一样，都是一种推测。

我强调这种比较并不意味着鼓励一种还原主义的方法论：这种方法论断言，我们只应该谈论从形式的观点来看能够完全实现的还原。没有什么可能比这离真理更遥远。作为一名顽固

守旧的多元论者，我很高兴尊重推测和结果，特别是考虑到对于任何先进的科学来说，把特定的结果建立在还原的基础上是多么的困难。我们恰好需要承认为什么推测。

作为表征的定义。在转向把一个理论模型的同构作为表征的一般形式进路之前，明确地考虑思考一个理论的概念表征的最重要的方式是重要的，这种方式并不是根据理论的模型。这条进路是在一个理论中用该理论的原始概念定义新概念的有用而普遍的实践。

更明确地说，一个理论中的第一个定义是一种特定形式的一个句子：即用该理论的原始符号确定它的一个新符号的含义。一个理论中的第二个定义也是一种特定形式的一个句子：即用该理论的原始符号和第一次定义的符号，确定它的第二个新符号的含义。接下来的定义依次类推。值得注意的是，一个理论中的这些定义每一次都以某个固定顺序引入一个定义。由于这个固定的顺序，我们总是可以有意义地谈论该理论中的*前面的*定义。通常便于接受的观点是，必须只用该理论的原始符号来定义任何一个被定义的符号。在这种情况下，根本不需要以某种固定的顺序引入定义。然而通常的数学实践是，用以前所定义的符号来定义新的符号；于是，需要有一个固定的定义顺序才能精确地解释这种实践。

从推理的逻辑观点来看，一个理论中的一个定义只被看成是一个新的公理或前提。但这并不意指一个定义以任何一种重要的方式巩固了该理论。引入一个新符号是为了促进对该理论结构的演绎研究，而不是附加到那种结构当中。使描述定义的这些直观思想变得更加清楚的两个标准是：(1) 一个已定义的符号应该总是能从该理论的任何一个公式中消除的；(2) 一个新的定义不允许证明以前无法证明的旧符号之间的关系；也就

是说,它没有起到一个创造性公理的作用。①

这种根据定义表征的一个平凡的但却普遍存在的例子 54 是,根据严格顺序$\succ$的弱不等式$\succeq$、等式$\approx$和反向不等式$\preceq$的定义。

根据定义表征的一个很成功例子是,从只有一个原始概念的集合项开始,在集合论中定义所有的标准数学概念。(关于这方面的进展,可参见 Suppes, 1960b)。

3.2 模型的同构

可能适用于一个理论的最一般而有用的集合论概念是一个理论的两个模型或结构是同构的概念。大体上说,从一个理论的基本概念的观点来看,当它的两个模型表现出相同的结构时,它们就是同构的。对一个特殊的理论来说,同构的形式定义的观点会使**相同结构**这个概念变得更精确。然而,我要强调的是,一个理论的两个模型的同构定义,不依赖于该理论的详细本性,而事实上是非常独立的,通常被称为是"与公理无关的"。运用"与公理无关"这个短语表明,同构定义只依赖于一个理论模型的集合论特征。因此,两个理论,它们的模型有相同的集合论特征,但它们的重要公理却是相当不同,它们将运用相同的

---

① 关于定义理论的详细阐述可参见 Suppes (1957/1999,第 8 章)。在同一章讨论的一个相关问题涉及,根据其他概念,一个理论中的一个原始概念的可能的可定义性问题。帕多阿原理(Padoa's principle)(1902)为根据该理论的其他概念来证明一个概念的独立性,即不可定义性,提供了一个明确而直观的方法。这个原理很简单:为了证明一个给定概念的独立性,给定该理论的公理的两个模型,使得给定的原始概念在这两个模型中是不同的,但其余的原始概念在这两个模型中都是相同的。

同构定义。①

正如第二章所定义的那样,通过给出代数的同构定义,可以使这些思想变得更加明确。这里,一个结构(A, $\circ$, e, $^{-1}$)是一个代数,如果 A 是一个非空集,$\circ$是从 A×A 到 A 的一种二元运算,e 是 A 的一个元素,以及$^{-1}$是从 A 到 A 的一种一元运算。

定义 1. 一个代数 $\mathfrak{A}=(A, \circ, e, ^{-1})$ 与一个代数 $\mathfrak{A}'=(A', \circ', e', ^{-1'})$是同构的,当且仅当,存在着一个函数 f,使得

(i) f 的定义域是 A 和 f 的值域是 A',

(ii) f 是一个一一对应的函数,

(iii) 如果 a 和 b 在 A 中,那么,$f(a \circ b) = f(a) \circ' f(b)$,

(iv) 如果 a 在 A 中,那么 $f(a^{-1}) = f(a)^{-1'}$,

(v) $f(e) = e'$。

55 当我们问自己,两个不同的对象是否有相同的结构时,我们显然是问关于对象所属的概念的那个集合的问题。一件容易的事情是表明,刚才定义的同构关系是代数之间的一种等价关系,即它是自反的、对称的和可传递的。作为一个相当有趣的例子,我们可以考虑测量理论中应用的不同的但同构的两个群。设一个群是整数的加法群。在这种情况下,集合 A 是所有整数的集合,运算$\circ$是加法运算,单位元素 e 是 0, 以及逆运算$^{-1}$是负运算。作为与第一个群同构的第二个群,考虑 2 的所有整数幂的乘法

① 同构的词源学的起源来自希腊的相同形式的含义。正如我在本章的最后一节(3.6)所论证的那样,亚里士多德还有柏拉图著作中的形式概念,非常接近于这里所用的结构概念。说同构的两种结构具有相同的形式,是很自然的。此外,坚持要准确地说清楚,一个实体的结构同构于另一个实体的结构,相当接近地对应于亚里士多德的相关的形式概念——房子是质料的形式,是由砖块建成的,但一块砖是使房子成形的质料的形式。从一种更加技术而抽象的观点来看,我们能说,在外延上,一种结构的形式,在这里所用的意义上,是同构于它的所有结构的类,在这里,我希望把这个类限制在某个给定的区域内,以免出现可能的悖论。

群。在这种情形中,集合 A'是等于 2 的某个整数幂的所有数的集合,$\circ'$是乘法运算,单位元素是整数 1,逆运算是标准的倒数运算,即 x 的倒数是 $1/x$。为了确定两个群 $\mathfrak{A}=(A,+,0,-)$ 和 $\mathfrak{A}'=(A',\cdot,1,^{-1})$ 是同构的,我们可以用函数 f,使得集合 A 中的每一个整数 n 都有

$$f(n)=2^n$$

那么,很容易查明,f 的值域是 A', f 是一一对应的,而且,

$$\begin{aligned} f(m\circ n) &= f(m+n)=2^{m+n}=2^m\cdot 2^n \\ &= f(m)\cdot f(n)=f(m)\circ' f(n) \end{aligned}$$

$$f(n^{-1})=f(-n)=2^{-n}=\frac{1}{2^n}=f(n)^{-1'}$$

和

$$f(0)=2^0=1$$

应该明显的是,如果我们设加法群的对象集是所有实数的(正的或负的)集合,乘法群的对象集是所有正的实数的集合,那么在加法群与乘法群之间可能有相同的同构。从测量理论的观点来看,这种同构是令人感兴趣的,主要因为它意味着没有在加法表征和乘法表征之间作出选择的数学基础。外延量的标准讨论,例如,关于测量质量或距离的那些讨论,通常并不强调一个乘法表征与一个加法表征一样是可接受的和正确的。因为质量或距离的测量绝对不是负的,有人可能会认为,这种评论恰好不适用于群,因为所有考虑的加法群都有负数作为群的元素。回答是,在考虑质量或距离的实际测量时,我们只限于所讨论的加法群的正元素的半群。不过,这种观点的详情与这里的评论无关。关于同构或结构的同一性的早期评论,只与一个特定的概念集

相关，注意，整数和 2 的幂的乘法群在许多数论特性上是完全不同的。

作为通过定义一个集合论谓词使理论公理化的另一个简单例子，我们可以考虑第二章讨论过的定序测量理论。这个理论
56 的模型习惯上被称为弱排序，而且我们将用这个术语来定义适当的谓词。这个理论模型的集合论结构是一个非空集 A 和在这个集合上定义的一种二元关系 R。设我们把这样一对 $\mathfrak{A} = (A, R)$ 称为一种简单的关系结构。在第二章定义 1 的风格中，我们就有下列定义。①

定义 2. 一个简单的关系结构 $\mathfrak{A} = (A, R)$ 是一个弱排序，当且仅当，对于 A 中的每一个 a、b 和 c 来说：

(i) 如果 aRb 和 bRc，那么 aRc；

(ii) aRb 或 bRa。

简单关系结构的这种同构定义应该是显而易见的，但不管怎样，为了明确起见，我给出了这种定义，而且，再一次强调，同构定义只依赖于简单关系结构的集合论结构，不依赖于所强加的任何重要的公理。

定义 3. 一个简单关系结构 $\mathfrak{A} = (A, R)$ 同构于一个简单关系结构 $\mathfrak{A}' = (A', R')$，当且仅当，存在一个函数 f，使得

(i) f 的定义域是 A，并且 f 的值域是 A'；

(ii) f 是一个一一对应的函数；

(iii) 如果 a 和 b 在 A 中，那么 aRb，当且仅当，$f(a)R'f(b)$。

为了举例说明这种同构定义，让我们考虑这样的问题："有

---

① 对于同构的分析来说，我在这里用了一般的二元关系符号"R"，而不用在第二章第二节提出的具有示意性的符号"$\geq$"，后面，在方便的地方，我也用"$\leq$"表示"$\geq$"的逆向。

相同元素个数的任何两个有穷弱排序,是同构的吗?”从直观上看,答案似乎显然应该是否定的,因为在其中的一个弱排序中,所有对象可能都具有相互关系 R,而在另一个弱排序中,却并非如此。根据我们能够构造的最小定义域的反例表明,不存在这样一种同构吗?很显然,两个单一元素的集合就做不到,因为在同构范围内,一个元素只有一种弱排序,即这种排序使得一个元素在给定关系 R 中就是自身。然而,增加一个元素就能找到一个反例。在一个弱排序中,我们能够设 R 是全域关系,也就是说,$R = A \times A$,即 A 与自身的笛卡儿积,在另一个弱排序中,设 R' 是满足一个弱排序公理的一种“最低限度的”关系。更加形式地说,设

$$A = \{1, 2\}$$

$$R = \{(1, 1), (2, 2), (1, 2), (2, 1)\}$$

$$A' = A$$

$$R' = \{(1, 1), (2, 2), (1, 2)\}$$

然后,很容易查明,$\mathfrak{A} = (A, R)$ 和 $\mathfrak{A}' = (A', R')$ 都是含有基数 2 的定义域的弱排序,但是,$\mathfrak{A}$ 不可能同构于 $\mathfrak{A}'$。因为假定有一个函数 f 来确定这样一种同构。那么我们会有

$$1R2 \quad \text{当且仅当} \quad f(1)R'f(2)$$

和

57

$$2R1 \quad \text{当且仅当} \quad f(2)R'f(1)$$

但我们也有 $1R2$ 和 $2R1$,由此

$$f(1)R'f(2) \text{ 和 } f(2)R'f(1) \tag{1}$$

但这是不可能的,因为如果 $f(1)=1$,那么 $f(2)=2$,因此,从

(1)有 $2R'1$,可是,我们并没有 $2R'1$。另一方面,作为其他可能的惟一一一对应函数时,如果 $f(1)=2$,那么 $f(2)=1$,再者,我们必须从(1)有 $2R'1$,与 R'的定义相反。

3.3 表征定理

在试图描绘一个理论模型的本性时,同构概念成为一个核心问题。也许,对一个理论模型的最好的和最有说服力的描述,是根据有意义的表征定理来表达的。正如前面非形式地概述的那样,一个理论的**表征定理**的含义如下:表明以某种直观上明确的概念推理而著称的一个理论模型的特定类,是为了在同构范围内举例说明该理论的每个模型。更精确地说,设 $\mathfrak{M}$ 是一个理论的所有模型的集合,再设 $\mathfrak{B}$ 是 $\mathfrak{M}$ 的某个特异子集(distinguished subset)。关于 $\mathfrak{B}$ 的 $\mathfrak{M}$ 的一个表征定理,将包含下列断言:即给定 $\mathfrak{M}$ 中的任何一个模型 M,在 $\mathfrak{B}$ 中都有一个同构于 M 的模型。换言之,从这种理论的观点来看,在限于集合 $\mathfrak{B}$ 的范围内,例示了模型的每一种可能变化。应该显而易见的是,一个平常的表征定理总是能通过接受 $\mathfrak{B}=\mathfrak{M}$ 得到证明。一个表征定理恰好与模型类 $\mathfrak{B}$ 的直观意义一样令人感兴趣,仅此而已。一个简单而漂亮的表征定理的例子是凯莱定理(Cayley's theorem):每个群都同构于一个变换群。正如它在19世纪所出现的那样,一个群的概念源于把一个集合映射到自身的一一对应函数的考虑。这样的函数通常被称为变换。有趣而令人惊讶的是,群的基本公理在这种抽象的意义上足以描述变换,也就是说,能够表明,在这种意义上,这些公理的任何一个模型,即任何一个群,都同构于一个变换群。(关于这个定理的讨论与证明,参见 Suppes, 1957/1999,第12章。)

表征定理的某些情形是特别有趣的。当集合 $\mathfrak{B}$ 是一个单元集合，即恰好只有一个元素的集合时，那么该理论被说成是范畴的。换一种方式说，当任意两个模型是同构的时，一个理论就是范畴的。因此，范畴论在同构范围内确实只有一个模型。当运用标准的集合概念时，范畴理论的例子是基本的数论和具有相同标准的集合概念的基本的实数理论。有时，有人断言，19世纪的数学与 20 世纪的数学之间的主要差别之一是，19 世纪的数学是研究范畴的数学理论，而 20 世纪的数学则是研究非范畴的理论。能从历史上作出这种区分，是令人怀疑的，但是在范畴理论与非范畴理论之间肯定有一个相当明确的概念区别。有一种明显的感觉是，非范畴的理论更加抽象。

58

从心理学的观点来看，有可能举出关于这种观点的一个好的例子是，当模型的类变得如此之大以致一种典型模型的任何一个简单的意象或图像都是不可能时，可以把一个理论看成是抽象的。模型的范围是五花八门的；这个理论完全是非范畴的。另外一种密切相关的“抽象”感是，正如在群的情形中一样，该理论的最初模型的某种直观的也许通常是复杂的特性，一直没有得到考虑，而且我们现在准备讨论满足一个理论的模型，即使这些模型比最初的直观模型有更简单的内在结构。

模型的同态。在纯数学的许多情况下，根据模型同构的表征定理原来还不如根据更弱的同态概念的表征定理令人感兴趣。在科学哲学中，这种类型的一个好的事例是由测量理论提供的，而且在这种语境中我们能够举例说明从同构到同态的概括。大体上说，当我们考虑一般的测量实践时，很显然，根据结构的同构概念，我们把同构看成是在测量理论的经验模型和数值模型之间确立的。我们所说的**经验模型**是指，在这个模型中，

基本集合是经验对象的集合,所说的数值模型是指,在这个模型中,基本集合是数的集合。然而,对这个问题的一个稍微更加详细的考察表明,关于同构的困难很快就产生了。在实在太多的测量情形中,不同的物理对象被赋予相同的数,因而,模型的同构所要求的一一对应关系不成立。幸运的是,为了对经验模型和数值模型之间关系的测量理论作出一种适当的说明,这是惟一要考虑的,在这种考虑中,我们必须改变一般概念。构思出一般的同态概念就是为了恰好与这种情境相符合。像前面所定义的那样,为了获得两个代数或两种简单关系结构的同态的形式定义,我们只需要放弃下列必要条件:确立这种同构的函数是一一对应的。当这种函数是多对一而不是一对一时,我们拥有了一种不是同构的同态。①

通过考虑作为一种测量理论的弱排序理论,可能会使这些评论变得更加具体。很容易给出两种弱排序的一个简单例子,使得第一种弱排序与第二种弱排序同态,但不同构。

设

$$A = \{1, 2\}$$

$$R = \{(1, 1)(2, 2)(1, 2)(2, 1)\}$$

$$A' = \{1\}$$

$$R' = \{(1, 1)\}$$

59 和

① 一个更弱的同态概念一般在代数中使用。例如,根据函数 f 是从 A 映射到 A',结构 (A, R) 和 (A', R') 是同态的就是,如果 xRy,那么 $f(x)R'f(y)$,而不是当且仅当。然而,在测量理论和科学哲学的其他应用中,这里所用的这种定义是更加令人满意的。

$$f(1) = 1$$

$$f(2) = 1$$

从这些定义来看,立刻显而易见的是,弱排序 $\mathfrak{A} = (A, R)$ 在函数 f 的条件下与弱排序 $\mathfrak{A}' = (A', R')$ 是同态的。这种观点是,我们有

$$1R2,\text{当且仅当},f(1)R'f(2)$$

以及

$$2R1,\text{当且仅当},f(2)R'f(1)$$

并且这两种等价性成立,只是因为

$$f(1) = f(2) = 1$$

另一方面,只在基数的基础上考虑,$\mathfrak{A}$ 也显然不同构于 $\mathfrak{A}'$,因为集合 A 有两个元素,集合 A'有一个元素。$\mathfrak{A}'$也明显地不同态于 $\mathfrak{A}$。这也是根据基数的考虑得出的,因为根本没有这样的函数:它的定义域是集合 A',它的值域是集合 A。正如这个例子所例示说明的那样,在一个理论的模型之间的同态关系并不是一种等价关系;它是自反的和传递的,但不是对称的。

我所说的一个**数值**的弱排序是指一个弱排序 $\mathfrak{A} = (A, \leqslant)$,这里,$A$ 是数的集合。在数值关系$\geqslant$同样能够被选中的意义上,一个弱排序中表示关系 R 的数值关系$\leqslant$的选择是任意的,不过,如果有第一个元素,却没有最后一个元素,那么用$\leqslant$稍微更加自然,自然数集合的情形也是如此。然而,选择两种关系$\leqslant$或$\geqslant$当中的一种在直观上是惟一合理的可能性。下列定理为有穷弱排序提供了一个同态的数值表征定理,因而,由于它的数值表征,使这种有穷弱排序理论成为一种测量理论。

定理 1. 每一个有穷弱排序都与数值的弱排序同态。[①]

60　定理 1 只限于有穷弱排序是有原因的；假如去掉这种限制，则是错误的。经典的反例是，词典式的平面序列。设 A 是实数的所有序对的集合，再设关系 $\preceq$ 是通过下列等价关系定义的：$(x_1, x_2) \preceq (y_1, y_2)$，当且仅当，$x_1 < y_1$，或者 $x_1 = y_1$ 和 $x_2 \leqslant y_2$，设想，存在着满足下列等价关系的一个真值函数 f：

$$f(x_1, x_2) \leqslant f(y_1, y_2)，当且仅当，(x_1, x_2) \preceq (y_1, y_2) \tag{1}$$

我们根据 $x_2 < y_2$ 来确定 x_2 和 y_2，并且，对于每个 x_1 来说，定义

$$f'(x_1) = f(x_1, x_2)$$

---

① **定理 1 的证明**：设 $\mathfrak{A} = (A, \preceq)$ 是一个有穷弱排序。也许，最简单的进路是，首先在 A 中形成对象的等价类，根据 $\preceq$，把相关的明显的等价关系 $\approx$ 定义为

$$a \approx b，当且仅当，a \preceq b \ \& \ b \preceq a$$

因此，用标准符号"$[a]$"表示等价类，即

$$[a] = \{b: b \in A \ \& \ a \approx b\}$$

我们首先按照 $\preceq$ 排列等价类。明确地说，我们定义

$$[a] \preceq^* [b] 当且仅当，a \preceq b$$

可以直接证明，在等价类的集合 $A/\approx$ 中，$\preceq^*$ 是自反的、反对称的、传递的和关联的，或换言之，它是 $A/\approx$ 的一种简单排序。根据假设，A 是一个有穷集，因此，$A/\approx$ 一定是有穷的。设在弱排序 $\preceq^*$ 条件下，$[a_1]$ 是 $A/\approx$ 的第一个元素，$[a_2]$ 是第二个元素……以及 $[a_n]$ 是最后一个元素。现在考虑在 $A/\approx$ 上定义数值函数 g，定义如下：

$$g([a_i]) = i，其中，i = 1, \cdots, n。$$

那么，g 在排序 $\mathfrak{A}/\approx = (A/\approx, \preceq^*)$ 和数值排序 $\mathfrak{N} = (N, \leqslant)$ 之间建立了一种同构，这里，N 是从 1 到 n 的正整数的集合。（这部分证明的细节是冗长的，但是明显的。）我们就在 A 上定义数值函数 f，对于在 A 中的每一个 b 来说，根据：

$$f(b) = i，当且仅当，b \in [a_i]。$$

即如果 b 是在排序 $\preceq^*$ 条件下的第 i 个等价类。函数 f 在 $\mathfrak{A}$ 和 $\mathfrak{N}$ 之间确立了所渴望的一种同态。

$$f''(x_1) = f(x_1, y_2)$$

根据这些函数,我们从实数到区间定义下列函数 g:

$$g(x_1) = [f'(x_1), f''(x_1)]$$

根据排序是词典式的这种假设,g 一定是一一对应的,因为两个不同的数被映射到两个不重叠的区间。例如,如果 $x_1 > x_1'$,那么 $f'(x_1) = f(x_1, x_2) > f(x_1', y_2) = f''(x_1')$。但众所周知,在不可数的实数集合和可数的非简并的不重叠区间之间根本不可能是一一对应的。因此,这样的函数 g 是根本不可能存在的,而且,对于词典式排序来说,更不可能存在满足(1)的函数 f。

前面对一阶理论的任意大基数的模型作出的评论也提供了一种证明:在定理 1 中,限于有穷弱排序是有必要的。在前面部分,我们为了确保一个无穷弱排序与一个数值的弱排序是同态的,在考虑必须为弱排序附加的那些公理时,显示了在集合论框架内阐述理论的一个优势。当时就讲过,足以确保这种结果的必要条件和充分条件,没有简单的一阶陈述。在一阶框架内很难阐述这种必要条件的原因是,需要量化这种弱排序的基本集合 A 的无穷子集。

为了陈述想要的定理,需要一个基本概念。设 $\mathfrak{A} = (A, \leq)$ 是一个简单的关系结构,再设 B 是 A 的一个子集。通过下面的等价关系,根据$\leq$,定义严格的排序关系$<$

$$a < b, \text{当且仅当}, a \leq b \text{ 和非 } b \leq a$$

然后,我们说,B 是在 A 中$\leq$的*稠密序列*,当且仅当,对于 A 中而不是 B 中的每个 a 和 b,使得 $a < b$,在 B 中存在一个 c,使得 $a \leq c$ 和 $c \leq b$。注意,有理数的可数集是在关于自然数排序$\leqslant$的所有实数的不可数集中的稠密序列。在可数的有理数和所有实数之间的这种关系恰好对于在一种无穷弱排序和一种数值的弱 61

排序之间一种同态的存在来说，是充分必要的。在应用可数性条件时，产生了一个不太复杂的问题，即稠密序列子集的元素不一定是等价的。这个附加条件在下面的定理陈述中变得更加明确。

定理 2. **设 $\mathfrak{A} = (A, \preceq)$ 是一个无穷弱排序。于是，$\mathfrak{A}$ 与一个数值的弱排序同态的充分必要条件是，存在着 A 的一个可数的子集 B，使得，(i) B 是 A 中 $\preceq$ 的稠密序列，(ii) B 中没有两个元素具有关系 $\approx$。**①

---

① **定理 2 的证明：** 康托尔(Cantor, 1895)把这个证明与对连续统的经典序数描述联系起来。我们在这里不提供所有的细节，只是概述主要的轮廓；关于某些附加的细节和相关定理，参见 Sierpenski(1958，第 11 章)。

为了证明这个条件的充分性，设 B 是有特性(i)和(ii)的一个可数的子集。此外，如果 A 相对于这种排序有两个端点，那么我们可以在不失一般性的前提下，在 B 中包括这两个端点。首先，我们知道，在 B 上有一个实值函数 f，确定 $\mathfrak{B} = (B, \preceq)$ 是与一个数值弱排序同态的，恰好是因为 B 是可数的(关于这种情况的证明，参见 Suppes and Zinnes, 1963, 定理 6)。现在，根据在 B 上的 $\preceq$ 的稠密序列条件，不在 B 中的 A 的每个元素 b 都定义了一个在 B 中的一种*划分*，也就是说，把 B 划分成两个集合，

$X = \{a: a \in B \ \& \ a \preceq b\}$ 和 $Z = \{c: c \in B \ \& \ b \preceq c\}$。设

$$r_1 = \underset{a \in X}{g.l.b.} \ f(a)$$

和

$$r_2 = \underset{c \in Z}{l.u.b.} \ f(c)$$

我们于是通过下列定义把 f 扩展到 b：

$$f(b) = \frac{r_1 + r_2}{2}$$

这很容易表明，函数 f 因此从 B 扩展到 A，确立了 $\mathfrak{A}$ 与一个数值排序的同态。例如，如果 a_1 和 a_2 在 A 中，而不是在 B 中，那么如果 $a_1 \preceq a_2$，就在 B 中存在着一个 c，使得 $a_1 \preceq c \preceq a_2$，因此，

$$f(a_1) \leqslant f(c) \leqslant f(a_2)$$

这部分论证的其余细节由读者来提供。 (转下页)

有人已经评论说，在一阶逻辑的范围内，根本没有直接的方式来阐述，正如定理 2 中所要求的那样，可数集 B 的存在性。也许，在科学哲学中，有一个更加明显但还没有得到充分认可的观点是，定理 2 的证明用了比初等逻辑更多的数学方法。限于在一阶逻辑的框架内讨论科学哲学中的理论问题，这种限制通常是意指，这种讨论可能达到的那种程度的一种限制，而不是为了达到当前的目的所特别构造的高度简化的科学理论。对理论进行集合论阐述的一个优势是，所有必要的数学工具都能直接地运用于该理论的阐述和它的模型结构的演绎分析。 62

模型的嵌入。我们已经明白，两个模型是同态的概念是对两个模型是同构的概念的概括。在两个模型之间还有一种更一般因而更弱的关系是，把一个模型嵌入到另一个模型中的关系。

（接上页）为了证明(i)和(ii)的必要性，我们设想，我们有一个函数 f 来确立 $\mathfrak{A}$ 与一种数值排序是同态的。由于有理数的可数性，含有有理数端点的实数的非空区间 I_i 的集合是一个可数的集合。我们下一步通过承认在每个区间 I_i 的 f 条件下的逆像，构造 A 的对应区间 J_i。从每个非空的 J_i，我们选择一个元素 a_i。既然区间集合是可数的，所以 A 的元素 a_i 的集合 X 是可数的。第二，设 Re 是实数 r 的集合，使得对于 A 中的某个 b 来说，$f(b) = r$，和

$$f(b) - \underset{a \in Y}{l.u.b.}\ f(a) > 0$$

在这里，

$$Y = \{a: a \in A \ \&\ a \leq b\}$$

因为实数集合 Re 定义了实数的一个非重叠区间的集合，所以它至多是可数的（当然，Re 可能是空的，而且在某些情况下也将是空的）。设 X' 是在 Re 的 f 条件下的逆像。于是，$X \cup X'$ 是可数的。为了表明 $X \cup X'$ 是 A 中的稠密序列，设 c_1 和 c_2 是 A 中而不是 $X \cup X'$ 中的两个元素，使得 $c_1 \prec c_2$。现在，如果在 c_1 和 c_2 之间没有 A 的元素，那么与假设相反，c_1 是在 X' 中。另一方面，如果在 c_1 和 c_2 之间有元素存在，那么，比如说，至少存在着一个元素 d，使得 $f(d)$ 位于含有有理数端点的一个区间内，这个有理数的端点嵌套在区间 $[f(c_1), f(c_2)]$ 中，因而 $X \cup X'$ 在 A 中是稠密序列。这就完成了这个证明。

证明一个理论的嵌入定理(embedding theorem)就是证明,存在着一个令人感兴趣的模型类 $\mathfrak{B}$,使得该理论的每个模型都与属于 $\mathfrak{B}$ 的一个子模型是同构的,或至少是同态的。子模型的精确定义从一个理论到另一个理论是稍有变化的,取决于它的模型的集合论特征。例如,如果 $\mathfrak{A} = (A, \circ, e, ^{-1})$ 是上面定义的一个代数,那么代数 $\mathfrak{A}' = (A', \circ', e', ^{-1'})$ 是 $\mathfrak{A}$ 的一个子代数,如果 A'是 A 的一个子集,$\circ'$是限于 A'中的运算$\circ$[即 $\circ' = \circ \cap (A' \times A')$, $e' = e$, 以及$^{-1'}$是限于A'中的运算$^{-1}$]。① 在简单关系结构的情况下,这个定义还更简单。设 $\mathfrak{A} = (A, R)$ 和 $\mathfrak{A}' = (A', R')$ 是两个这样的结构。于是,如果 A'是 A 的一个子集,而且,R'是限于A'中的关系R,即 $R' = R \cap (A' \times A')$,那么 $\mathfrak{A}'$是 $\mathfrak{A}$ 的一个子模型。

沿着下列思路能够把定理 1 阐述为一个嵌入定理。设 Re 是实数集。于是,马上可以看出,像前面定义的那样,$(Re, \leqslant)$ 是一个数值弱排序,而且每一个有穷弱排序都能在同态的意义上被嵌入到$(Re, \leqslant)$中,即都与$(Re, \leqslant)$的子模型是同态的。

在这里,我结束了对模型和理论的一般讨论。② 本章的后面两节专门列举两个表征定理的例子,第一个是基本的测量理

---

① 原文在这里缺少了一个右括号符号,与作者沟通后,在这里加上。——译者

② 我在这个脚注中确实补充了一个最后的一般评论,表征定理的语言,除了用特殊学科熟悉的惯用语替代之外,通常并没有使同构的本性更明显。下面是我意指的两个例子。

(i) "一个粒子系统的质心的运动像一个粒子的运动一样,这个粒子具有的质量与这个系统的质量相等……"代替了"一个粒子系统的质心在它的运动中是与一个粒子系统的运动同构的,这个粒子具有质量……"(7.1 节的定理 4)

(ii) "四个可观察量满足贝尔不等式,当且仅当,它们有一个联合概率分布",代替了"给定四个可观察量的六对分布,能够在同构的意义上被嵌入到四个可观察量的一个普遍的联合分布当中"。(7.2 节的定理 6)

论，另一个是通过抽象的通用计算机的一个程序对任何一个可计算函数的重要表征。再接下来的一节，即本章的最后一节，返回到更一般的论题，特别是，有关心理表征本性尤其是关于同构问题的观点。 63

3.4 基本测量结构的表征

在这一节，我考虑测量结构的最简单的但非无足轻重的几个例子。对象或刺激物的基本集合在所有情况下都是有穷的，而且不同结构的基本公理的适当性非常依赖于这种有穷性。除了它们的有穷性之外，所考虑的结构的典型特征是，在适当的意义上，这些对象沿着连续统等区间隔开，比如，所测量的特性。有穷性和等区间的限制极大地简化了测量的数学，但幸运的是，不是下列情形：简化伴随着与实际经验的应用完全分离开来。有穷性和等区间是许多标准尺度的典型属性，例如，普通的直尺，在实验室或商店里所用的天平的标准秤砣，或测量压强、温度或体积所熟悉的几乎任何一种计量器。

我研究四种这样的结构及其表征。其中每一种结构都对应于克兰茨(Krantz)、卢斯(Luce)、苏佩斯和特韦尔斯基(Tversky)的综合性论文中所分析的更一般的结构集合。这四种结构是，外延测量(extensive measurement)、差测量(difference measurement)、对分测量(bisection measurement)和联合测量(conjoint measurement)。这里给出的这种分析来自苏佩斯(Suppes, 1972)。

外延测量。在科学史与哲学史上，外延属性或量与内涵属性或量之间的区分是很古老的。外延量是能够相加的量；例如，质量和长度是外延量或参量。相反，内涵量是不可能相加的，即

使它们能够被测量。比如,具有相同温度的两个容量的气体,不能结合在一起形成具有两倍温度的一种气体。有些理论家,例如,坎贝尔(Campbell, 1920, 1928)曾主张,内涵量的基本测量是不可能的。坎贝尔等人的否定论证根本没有说服力,而且测量结构的许多例子,包括本节后面提供的那些例子,对坎贝尔的论点提供了具体的反驳。

我在这一节根据想到的三种特殊解释提出了外延测量的公理。一个是用天平测量质量,另一个是用固定标尺测量长度,还有一个是测度主观概率。① 其他解释肯定是可能的,但我将限于详细地评论这三种解释。

从形式的观点来看,这种基本结构是三元组$(\Omega, \mathcal{F}, \succeq)$,这里$\Omega$是一个非空集,$\mathcal{F}$是$\Omega$的一个子集族,关系$\succeq$是在$\mathcal{F}$上的一种二元关系。通过用$\Omega$的子集作为对象,我们避免了需要一个不同的基本的并置概念。作为一般的结构条件,将要求$\mathcal{F}$是在Ω上的一个**集合代数**,这恰好要求在集合的并集和补集条件下,$\mathcal{F}$是非空的和闭合的,即如果A和B在$\mathcal{F}$中,那么$A \cup B$和$-A$也在$\mathcal{F}$中。

64

对于所提到的这三种情况来说,原始概念的预期解释是相当明显的。在质量的情况下,Ω是物理对象的一个集合,而且对于两个子集A和B来说,$A \succeq B$,当且仅当,断定对象集合A至少与集合B一样重。也许值得强调的是,天平的几种不同用法适合于达成一种比较判断。例如,如果$A = \{a, b\}$和$B = \{a, c\}$,那么实际上不可能把A放在天平的一个秤盘里,同时,把B放在另一个秤盘里,因为对象a是两个集合中的成员。但

① 由于在第5章中广泛地使用了这种概率解释,下面,我把代表一个概率空间的可能结果集合的标准符号"Ω"用作本节外延测量的一般概念。

是,我们至少可以以两种不同的方式来进行这种比较。一种是只比较两个集合的不重叠的部分,在当前的情况下,恰好是比较$\{b\}$和$\{c\}$。一个相当不同的经验程序甚至不需要天平两端保持平衡:先在天平的另一个盘里放上沙石(也可能是水;但在每一种情况下,沙石或水都在小容器里)与 A 达到平衡,然后把 B 与所固定的沙石量进行比较。给定交集、并集和补集的集合论运算的标准含义,就不需要对这些运算,甚至是作为并置运算的集合的并集,作出额外的解释。

在固定标尺的情况下,集合 Ω 恰好是各种标尺的集合,而且,$A \geq B$,当且仅当,把集合 A 中的标尺一个挨一个地放在一条直线上,把集合 B 中标尺也同样排放,集合 A 的标尺形成的直线比集合 B 的标尺形成的直线要长。我们能轻易地给出如何准确地对长度进行这种定性比较的偏差。

在主观概率或客观倾向性的情况下,集合 Ω 是所考虑的实验或经验情境的可能结果的集合。在通常的概率概念的意义上,Ω 在 $\mathcal{F}$ 中的子集恰好是诸事件,而且,$A \geq B$,当且仅当,断定 A 至少与 B 是一样可能的。

由于受到有穷性和等区间两种限制,外延测量公理是在下列定义中给出的。在这个定义中和后面的定义中,我们根据一个弱排序,还有一个严格的排序,运用等价关系$\approx$的标准定义。这个定义恰好是:$A \approx B$,当且仅当,$A \geq B$ 和 $B \geq A$;$A \succ B$,当且仅当,$A \geq B$ 和非 $B \geq A$。

定义 1. 一个结构 $\Omega = (\Omega, \mathcal{F}, \geq)$ 是一个有穷的、等区间的外延结构,当且仅当,Ω 是一个有穷集,$\mathcal{F}$ 是在 Ω 上的一个集合代数,而且 $\mathcal{F}$ 中的每个 A、B 和 C 都满足下列公理:

1. 关系$\geq$是 $\mathcal{F}$ 的一个弱排序;

2. 如果 $A \cap C = \varnothing$ 且 $B \cap C = \varnothing$,那么 $A \geq B$,当且仅

当，$A \cup C \succeq B \cup C$；

3. $A \succeq \varnothing$；

4. 非 $\varnothing \succeq \Omega$；

5. 如果 $A \succeq B$，那么在 $\mathcal{F}$ 中存在着的一个 C，使得 $A \approx B \cup C$。

从测量质量或长度的标准思想的观点来看，断言如果 $A \neq \varnothing$，那么 $A \succ \varnothing$，强化公理 3 是很自然的，但因为这并不是表征定理所要求的，也过分地限制在概率的情况下，所以公理越不充分，似乎越适当。

65

在阐述表征定理时，我们使用了从 $\mathcal{F}$ 到实数的**加法测量**概念 μ，即一个函数 μ 使得对于 $\mathcal{F}$ 中的任何一个 A 和 B：

i. $\mu(\varnothing) = 0$；

ii. $\mu(A) \geqslant 0$；

iii. 如果 $A \cap B = \varnothing$，那么 $\mu(A \cup B) = \mu(A) + \mu(B)$。

定理 1. *设 $\Omega = (\Omega, \mathcal{F}, \succeq)$ 是一个有穷的、等区间的外延结构。于是，存在着一种加法测量 μ，使得对于 $\mathcal{F}$ 中的每个 A 和 B，*

$$\mu(A) \geqslant \mu(B)，当且仅当，A \succeq B$$

此外，所有的单一元素集合最多属于 $\mathcal{F}$ 中的两个等价类；

如果存在着两个等价类，那么这两个等价类之一包括空集。

这个定理的证明和上述的一些讨论将会在苏佩斯的著作(1969a, pp. 4－8)中找到。

重要的是指出，对于在定义 1 中描述的外延测量理论中用到的原始概念的所有三种经验解释来说，每一种解释所要求的许多重要的经验程序目前缺失。如果你愿意的话，你可以向基础工程或物理学实验室里的学生询问用天平测量质量或用固定

标尺或卷尺测量长度的细节。但即使是这些细节也不足以作为对基本测量的一种说明。因为在更早的阶段,必须用这样的测量来确认这些学生所用的秤砣或标尺是有效的。例如,经验程序更详尽地阐述了创造一个标准秤砣的集合。在创造“相等”的标准秤砣时,或者在查明一个给定的标准秤砣的集合似乎是令人满意的时,定义 1 的公理开始发挥作用。我故意说“似乎”,因为在物理学或其他学科中,标准的测量程序或测试设备的经验研究很自然是无止境的。创造一种好的测量器械的艺术是一项复杂的和精致的艺术,当然,因此而居于许多学科的最前沿。我已经重复地论证过,这样的艺术不可能在小册子中得到充分地描述,它要求学徒式的训练其中的许多成分并不是用语言可以表达的。例如,请你想象一下,只凭读一本关于如何打网球的书或听几节讲座,一个人就能成为一名好的网球手吗? 任何一个实验室或为了制造出好的测量仪器所要求的一组制造技能,也是如此。①

这一节提出的不同形式类型的基本测量理论,是测量的一个重要的形式问题,但决没有提供全部基础,因为从目前为止进行的讨论来看,这是显然的,我相信,以一种明确的方式作出完全阐述,是不可能的。但是,所停止的地方正好在形式上是不明确的。我们能够制造一个聪明的机器人,然后编程,让机器人从事生产很好地近似满足定义 1 的公理的标准测量仪的工作。如 66 果通过追踪每一步构造方式的标准方法来教机器人该怎么做,我们当然也能够使为机器人提供的程序形式化,即使我对这种观点持有保留态度。但是,抛开这种反对意见不说,设计与建造

---

① 甚至笛卡儿也承认这一点,正如在他的《光学》(1967/2001, pp. 162 - 173)的第十个谈话中,在描述如何切割透镜(即制造透镜)时,所能看到的那样。

机器人所要求的经验和学习，超越了任何一本语言操作手册。在我过去关于数据模型的一篇文章中(Suppes, 1962)，可以找到关于这些问题的更详细的评论。也参见 1.2 节中关于操作定义的讨论。

差测量。在谈到较早讨论的外延属性与内涵属性之间的区别时，我能轻易地举出一个例子支持把这一小节取名为**内涵测量**，因为差测量的特征是没有出现与加法相对应的运算，而且也没有为差结构设定对象或刺激物的有意义的结合。

如前所述，对象或刺激物的基本集合将是非空的和有穷的，但在差结构的情况下，集合中的关系将是一种四元关系。我用 A 表示对象的基本集合，用$\succeq$表示这种四元关系。支持四元关系$\succeq$的思想是，只有当 a 和 b 之间的定性的(代数的)差，至少与 c 和 d 之差一样大时，才有 $ab \succeq cd$。例如，在相似性判断的情况下，当受试者在一个心理学实验中判断，a 和 b 之间的相似之处至少与 c 和 d 之间的相似之处一样多时，关系$\succeq$才会成立，但应适当考虑这种差的代数符号。这种代数差要求在解释中照顾到；例如，在许多相似性实验中，一个自然的代数符号并没有与这种相似性联系在一起。满足当前要求的例子是对声音的效用或强度或音调的判断；事实上，在任何一种判断中，受试者都将承认和接受，这些判断很自然是在一维连续统中进行的。

我们恰好把这种四元关系$\succeq$定义为一种二元关系，$\succ$和$\approx$：

$$ab \succ cd\text{，当且仅当，非 } cd \succeq ab\text{，}$$

$$ab \approx cd\text{，当且仅当，}ab \succeq cd \text{ 和 } cd \succeq ab$$

同样方便的是，手头就有一些严格优先或偏好的二元关系的基本定义和中立的或不可区分的关系$\approx$。这些定义如下：

定义 2. $a \succ b$，当且仅当，$ab \succ aa$。

定义 3. $a \approx b$，当且仅当，$ab \approx ba$。
为了表达我们假设的等区间部分，我们需要补充一个定义，即这个定义要求在这种排序中相邻的对象是等区间的。为了达到这个目的，我们引入二元关系 J 的定义。这种二元关系恰好是直接前趋(immediate predcessor)关系。

定义 4. aJb，当且仅当，$a > b$，而且对于 A 中的所有 c 来说，如果 $a > c$，那么，或者 $b \approx c$，或者 $b > c$。
我现在转向有穷等差结构的定义。下面提供的那些公理是由苏佩斯和齐内斯(Suppes and Zinnes, 1963)给出的。

定义 5. 一个四元结构 $\mathfrak{A} = (A, \succeq)$ 是一个有穷的、等区间的差结构，当且仅当，如果 A 是一个有穷集，而且 A 中的每个 a、b、c 和 d 都满足下列公理： 67

1. 关系 $\succeq$ 是 $A \times A$ 的一个弱排序；
2. 如果 $ab \succeq cd$，那么 $ac \succeq bd$；
3. 如果 $ab \succeq cd$，那么 $dc \succeq ba$；
4. 如果 aJb 和 cJd，那么 $ab \approx cd$。

记得已经提到的经验解释，很容易掌握每个公理的直观解释。第一个公理只是要求这种四元关系 $\succeq$ 是根据对象或刺激物之间的定性的差别进行的一个弱排序。公理 2 在许多方面是最有说服力的和基本的公理。它表达了关系 $\succeq$ 的预期解释的一种简单的必要属性。公理 3 恰好表达了关于差的一种必要代数的事实。请注意，公理 1—3 是必要公理。只有公理 4 是充分的，而不是必要的；它把 J 与四元关系 $\approx$ 联系起来。这个公理的直观思想是，如果 a 位于 J 和 b 的关系中，以及 c 位于 J 和 d 的关系中，那么断定 a 和 b 之差与 c 和 d 之差是一样的，但应适当考虑代数符号。

根据这四个公理，我们能够证明下列表征定理。

定理 2. 设 $\mathfrak{A}=(A, \succeq)$ 是一个有穷的、等区间的差结构。那么存在着一个在 A 上的实值函数 φ，使得对于 A 中的每个 a、b、c 和 d，

$$\varphi(a)-\varphi(b)\geq\varphi(c)-\varphi(d)\text{，当且仅当，}ab\succeq cd\text{。}$$

这个定理的证明在本节的最后给出。此外，在通向证明这个定理的一系列基本的预备定理中，许多基本属性得到了条理化。

有人根据草率的检查推测，定义 5 的前三个公理描述了能为其找到一个数值表征的所有的有穷差结构。然而，斯科特和苏佩斯(1958)曾表明，所有可表征的有穷差结构的理论都不是通过这三个公理来描述的，而且确实不可能通过这些公理的任意简单的有限清单得到描述。

另一方面，有人认为，外加不必要的公理 4，更难满足这些公理，因为刺激物或对象的任何一个集合都不满足这些公理。然而，如果所研究的刺激物位于一个连续统中，那么将有可能选择一个满足这些公理的标准序列，正如在为使用天平选择标准秤砣集合情形中所做的那样。

对分测量。与差结构密切相关的关系结构是对分结构(bisection structure) $\mathfrak{A}=(A, B)$，这里，B 是在具有下列解释的有穷集合 A 上的一种三元关系：即 $B(a, b, c)$，当且仅当，b 是区间 a 和 c 之间的中点。对分方法在心理物理学中的历史很长久，但重要的是强调，满足以下给定的公理不要求假设基本的物理测量。我们大家需要的是一个定性连续统的直观思想，而为了达到形式的目的，甚至不需要这种思想。根据对分方法进行了基本的心理测试之后，希望这时有可能在可计算意义上找到把相同大小的物理测量与心理测试联系起来的简单的心理物

68

理函数。

以下给出的对分方法的公理意味着，在断言存在着一个数值表征函数之前，许多检查应该是令人满意的，但是，通常在报告使用对分方法的实验文献中，这些检查被忽略了。对于公理和定义的最简单的集合来说，我们把对分关系 B 和排序关系 $\geq$ 都看成是原始的，但很容易通过定义来排除 $\geq$。正如前面所定义的那样（定义 4），我们使用了二元关系 J。

定义 6. 一个结构 $\mathfrak{A}=(A, \geq, B)$ 是一个有穷的、等区间的对分结构，当且仅当，如果集合 A 是有穷的，而且 A 中的每个 a、a'、b、c 和 c' 都满足下列公理：

1. 关系 $\geq$ 是 A 的一个弱排序；
2. 如果 $B(abc)$ 且 $B(abc')$，那么 $c \approx c'$；
3. 如果 $B(abc)$ 且 $B(a'bc)$，那么 $a \approx a'$；
4. 如果 $B(abc)$，那么 $a > b$ 且 $b > c$；
5. 如果 aJb 且 bJc，那么 $B(abc)$；
6. 如果 $B(abc)$ 且 $a'Ja$ 和 cJc'，那么 $B(a'bc')$。

这些公理的直观解释是相对明晰的。第一个公理是已经熟悉的。公理 2 和 3 要求达到等价的两个端点的惟一性，这显然把对分从中间状态中分离出来。公理 4 以一种自然的方式叙述了三元对分关系和二元排序关系，尽管为对分关系强加了一种形式的约束，但这种约束通常会被省略。把这种有序属性增加到关系 B 中，简化了这些公理。公理 5 是等区间的一种强假设，以及公理 6 表达了这种等区间的一个附加特征。从前面给出的差结构的公理观点来看，有点令人吃惊的是，能够表明公理 6 独立于公理 5，但为了表明情况确实如此，很容易给出公理 1－5 的一个模型。因为根据：

$$B(abc)\text{，当且仅当，}aJb\text{ 且 }bJc$$

我们能接受一个模型和满足前五个公理。

这种表征定理假定了下列形式。

定理 3. 设 $\mathfrak{A} = (A, \succeq, B)$ 是一个有穷的、等区间的对分结构。于是，存在着一个在 A 上定义的实值函数 φ，使得对于 A 中的每个 a、b 和 c：

(i) $\varphi(a) \geqslant \varphi(b)$，当且仅当，$a \succeq b$

(ii) $2\varphi(b) = \varphi(a) + \varphi(c)$ 且 $\varphi(a) > \varphi(b) > \varphi(c)$，当且仅当，$B(a, b, c)$

这个定理的证明在本节的最后给出。

69 **联合测量**。在多种实验或观察环境中，已经证明是这种情形：测量一个量或特性是不可行的或在理论上是无趣的。令人感兴趣的是，对几个特性同时进行联合测量。在这种情况下，我们考虑相加的联合测量的公理。预期的表征是，我们使用了对象或刺激物的有序对。这些有序对的第一项来自一个集合，相应地代表了一种特性或一个量。这些有序对的第二项来自第二个集合的对象，代表了一个不同的量或特性。已知这种有序对的结构，我们只要求判断，一个联合的有序对是否比第二个有序对有更多的“联合”属性。

对有序对的这种考虑方式是自然的，很容易举例说明这种解释。假定要求我们判断一个人在一个机构中担任领导职务的潜力。我们有关于个人的根据顺序量表对他们的技术知识的评价和根据顺序量表测试感召力的评价。因此，就每个人而言，我们都能说，根据每一个量表，如何把他与其他任何一个人作出比较。问题是，作出的判断被看成是根据他们的整体能力在个人之间进行的。以下给定的公理表明，这种进一步的联合顺序条件足以确保一种有穷的等区间的联合测量，在这里的情况下，每一维都是等区间的。

作为第二个例子,一个对(a, p)能够代表具有强度 a 和频率 p 的一个音调,问题是判断两个音调中哪个听起来更响亮。因此,受试者判断 $(a, p) \succeq (b, q)$,当且仅当,音调(a, p)似乎至少与(b, q)一样高。容易给出的其他例子来自与经济学和物理学一样完全独立的那些学科,而且这些例子已经在克兰茨、卢斯、苏佩斯和特韦尔斯基的著作中(1971,第 6 章)进行了相当详细的讨论。

值得强调的是,本节所寻找的相加表征是一种特殊情况。可加性的概括在刚才引证的参考文献中进行了讨论。也应该注意到,本节限于有序对,而不是有序的 n 元组,这点不是本质的。

在转向(相加的)联合测量的公理之前,我们需要两个基本的定义,这两个定义允许我们根据个别分量来定义排序关系。在建立在有序对之间的排序关系的公理之基础上,我们将能证明,关于这些分量的排序关系也是弱排序。在下面的基本定义中,A_1 是第一个分量的集合,A_2 是第二个分量的集合。因此,当参考一个有序对(a, p)时,人们已经知道了,a 是在 A_1 中,p 是在 A_2 中。

定义 7. $a \succeq b$,当且仅当,对于在 A_2 中的所有 p,$(a, p) \succeq (b, p)$。

根据这种关系,我们以通常方式来定义 $a \succ b$ 和 $a \approx b$。同样,第二个分量也需要一个类似的定义。

定义 8. $p \succeq q$,当且仅当,对于 A_1 中的所有 a,$(a, p) \succeq (a, q)$。

我们也用了已经提出的在 $A_1 \times A_2$ 上的关系符号 $\succeq$,即 $(a, p) \succ (b, q)$,当且仅当,非 $(b, q) \succeq (a, p)$,以及 $(a, p) \approx (b, q)$,当且仅当,$(a, p) \succeq (b, q)$ 和 $(b, q) \succeq (a, p)$。在有穷等区间情况下的这个相加的联合测量公理在下列定义中体现

出来。

70 **定义 9.** 一个结构$(A_1, A_2, \geq)$是一个有穷的、等区间的相加联合结构，当且仅当，集合 A_1 和 A_2 是有穷的，并且 A_1 中的每个 a 和 b 以及 A_2 中的每个 p 和 q 都满足下列公理：

1. 关系$\geq$是在 $A_1 \times A_2$ 上的一个弱排序；
2. 如果 $(a, p) \geq (b, p)$，那么 $(a, q) \geq (b, q)$；
3. 如果 $(a, p) \geq (a, q)$，那么 $(b, p) \geq (b, q)$；
4. 如果 aJb 且(a, q)，那么 $(a, q) \approx (b, p)$。

定义 9 的四个公理的直观内容是明显的，但需要进行某种讨论。公理 1 当然是熟悉的一个弱排序的必要条件。公理 2 和 3 表达了一个分量独立于另一个分量的条件。这样，公理 2 说，如果对(a, p)至少与对(b, p)一样大，那么当 p 被 A_2 中的任何另一项 q 取代时，同样的关系还是成立的，公理 3 说，关于第二个分量也是如此。公理 4 是充分的但不是必要的。对于有穷等区间的差结构来说，公理 4 陈述了等区间假设，并几乎对应于相应的公理。

可以认为，单一性假设：如果 $(a, p) \approx (b, q)$ 和 $a \succ b$，那么 $q \succ p$，也需要被假定为是一个公理，但是，正如我们在证明表征定理时所表明的那样，这个附加假设是不必要的：只从前四个公理就能证明这一点。

我们现在转向的表征定理的陈述，正是假定了这种预期的形式。值得注意的惟一问题是，关于每个分量的两个实值函数通过相同的单位结合在一起。单位 α 在这个定理中的正常变化反映了这一点，但正如第 4 章(定理 4.5.4)所表明的那样，允许有不同的原点。

定理 4. 设 $(A_1, A_2, \geq)(A_1, A_2, \geq)$ 是一个有穷的、等区间的相加联合结构。于是，分别在 A_1 和 A_2 上存在着实值函

数 φ_1 和 φ_2，使得对于 A_1 中的 a 和 b 以及 A_2 中的 p 和 q，

$\varphi_1(a)+\varphi_2(q)\geqslant\varphi_1(b)+\varphi_2(p)$，当且仅当，$(a, q)\succeq(b, p)$

定理 2－4 的证明。①

定理 2 的证明。尽管下面的基本预备定理对提供定理 2 的证明不是必要的，但是它们在一种完全明确的讨论中是必需的，而且它们的结论在使读者有条理地思考不同结构(它们不像外延结构那么眼熟)时，也许是有用的。只在一些事例中给出这些预备定理的证明迹象。

所有的预备定理都涉及一种固定的四元结构 $\mathfrak{A}=(A, \succeq)$，以及前面定义的二元关系$\succ$、$\approx$和 J。

预备定理 1. 关系$\succ$在 A 上是非对称的和可传递的。

预备定理 2. 关系$\approx$在 A 上是自反的、对称的和可传递的。 71

预备定理 3. 对于 A 中的每个 a 和 b，下列之一是完全成立的：$a\succ b$、$b\succ a$ 和 $a\approx b$。

预备定理 4. 如果 aJ^nb，那么 $a\succ b$。(在这个预备定理和下列大多数预备定理中，这些证明要求运用对 n 的归纳。)②

预备定理 5. 如果 $a\succ b$，那么存在着一个(正整数)n，使得 aJ^nb。

预备定理 6. 如果 aJ^nb 且 aJ^nc，那么 $b\approx c$。

预备定理 7. 如果 aJ^mb 且 bJ^nc，那么 $aJ^{m+n}c$。

预备定理 8. 如果 aJ^mb 且 $aJ^{m+n}c$，那么 bJ^nc。

预备定理 9. 如果 $aJ^{m+n}b$，那么在 A 中存在着一个 c，使得

---

① 可以在不失连续性的条件下省略掉本节其余部分给出的证明。

② 在这个预备定理和后面的预备定理中，还有在定理 2 和后面定理的证明中，分别用了二元关系 J 的 n 次幂的概念。这个概念是递归定义的：aJ^1b，当且仅当，aJb；aJ^nb，当且仅当，存在着一个 c，使得 $aJ^{n-1}c$ 且 cJb。

$aJ^m c$。

预备定理 10. 如果 $aJ^n b$ 且 $cJ^n d$,那么 $ab \approx cd$。

预备定理 11. 如果 $ab \approx cd$,那么要么存在某个 n,使得 $aJ^n b$ 和 $cJ^n d$;要么存在着某个 n,使得 $bJ^n a$ 和 $dJ^n c$,或者 $a \approx b$ 和 $c \approx d$。

我们现在转向概述定理 2 的证明。设 c^* 是关于排序$\succ$的 A 的第一个元素。① 对于 A 中的每个 a,在 A 上的数值函数 φ 的定义如下:②

$$\varphi(a) = \begin{cases} 1 & \text{如果 } a \approx c^* \\ -n+1 & \text{如果 } c^* J^n a \end{cases}$$

然后,我们用这些基本的预备定理可以证明:

(i) $\varphi(a) > \varphi(b)$, 当且仅当, $a \succ b$;

(ii) $\varphi(a) - \varphi(b) \geqslant \varphi(d) - \varphi(e)$, 当且仅当, $ab \succeq de$。

定理 3 的证明。 我们从两个预备定理的证明开始。第一个对应于证明定理 2 时的预备定理 10,第二个对应于预备定理 11。应该注意到,恰好关于关系$\succ$和 J 的定理 2 的预备定理,在这里也是适用的。

预备定理 12. 如果 $aJ^n b$ 和 $bJ^n c$,那么 $B(abc)$。

证明: 我们进行归纳证明。对于 $n=1$ 来说,我们有公理 5。现在假定,我的归纳假设是有效的,并且,我们有

---

① 正如定义 $\varphi(a)$表明的,c^* 作为第一个元素没有惟一性。因此,最好说"设 c^* 是 A 的第一个元素。"这样,在排序中与 c^* 等价的 A 中的所有项都是第一个元素。用更技术性的语言来说,它们是等价类 $[c^*] = \{a: a \approx c^* \ \& \ a \in A\}$ 的项。

② 正如 $\varphi(a)$的定义所表明的,整个这一节实际上更自然地用了$\preceq$,而不是$\succeq$,以致 c^* 是这个序列中最小的元素,而不是最大的元素,那么这时 $c^* J^n a$, $\varphi(a) = n+1$。然而,我这里用了$\succeq$,因为我们在论述测量时用到了它(Krantz et al. 1971)。

(1) $aJ^{n+1}b$ 和 $bJ^{n+1}c$。

然后,我们从 J 的特性立即知道,在 A 中存在着元素 a' 和 c',使得 72

(2) aJa' 且 $a'J^nb$,

(3) bJ^nc' 且 $c'Jc$。

由此从(2)和(3)根据归纳假设得出

(4) $B(a'bc')$。

然后,再一次从(2)和(3),还有(4)和公理 6,我们像所渴望的那样推出

$$B(abc)$$

预备定理 13. 如果 $B(abc)$,那么存在着一个 n,使得 aJ^nb 且 bJ^nc。

证明: 从这个定理的假设和公理 4,我们有

$$a > b \text{ 和 } b > c$$

由此根据 J 的熟悉特性,存在着 m 和 n,使得

$$aJ^mb \text{ 和 } bJ^nc$$

现在假定 $m \neq n$; 为了确定性和不失一般性,我们可以假定 $m < n$。于是,存在着 d,使得

$$bJ^md$$

由此根据预备定理 1

$$B(abd)$$

但通过假设 $B(abc)$,由此根据公理 2

$$c \approx d$$

但另一方面,我们有

$$bJ^m c \text{ 和 } bJ^n c$$

这是不可能的,因此,我们像所希望那样得出结论 $m = n$。

给定预备定理 1 和 2,函数 φ 的存在性的证明,使得

(i) $\varphi(a) > \varphi(b)$, 当且仅当, $a > b$,

而且,

(ii) $\varphi(b) = \frac{1}{2}(\varphi(a) + \varphi(c))$ 和 $\varphi(a) > \varphi(b) > \varphi(c)$, 当且仅当,$B(a, b, c)$。

类似于定理 2 相对应部分的证明,也不需要详细地阐述。

定理 4 的证明。首先,在定义 9 的公理 1~3 的基础上,很容易证明,由两个分量 A_1 和 A_2 组成的序列的下列基本预备定理。

预备定理 14. 在 $A_i(i = 1, 2)$ 上的关系≈是一种等价关系,即它是自反的、对称的和可传递的。

预备定理 15. 在 $A_i(i = 1, 2)$ 上的关系>是对称的和可传递的。

73 **预备定理 16.** 对于 A_1 中的 a 和 b,正好下列之一是真的: $a \approx b$、$a > b$、$b > a$。对于 A_2 中的 p 和 q,恰好下列之一是真的: $p \approx q$、$p > q$、$q > p$。

我们下一步证明前面在讨论定义 9 的公理时提到的两个预备定理。

预备定理 17. 如果$(a, p) \approx (b, q)$ 和 $a > b$, 那么 $q > p$。

证明: 假定不是 $q > p$ 的情况。那么,通过预备定理 3,要么 $p \approx q$,要么 $p > q$ 。如果 $p \approx q$,那么 $(a, p) \approx (a, q)$,由此从可传递性和这个预备定理的假设推出,$(b, q) \approx (a, q)$,因此,$b \approx a$,这与预备定理 3 和假设 $a > b$ 相矛盾。另一方面,

从其他可替代的假定,即 $p > q$,也得出一种矛盾。因为我们有 $(a, p) > (a, q)$,由此从弱排序的熟悉特性和这个预备定理的假设推出 $(b, q) > (a, q)$,因此, $b > a$,这再一次与预备定理 3 和假设 $a > b$ 相矛盾。因此,我们得出的结论是,从这个预备定理的假设得出了所希望的 $q > p$ 。

预备定理 18. *如果* $(a, p) \approx (b, q)$ *和* $p > q$ *,那么* $b > a$。

证明: 在结构上与预备定理 4 完全一致。

我们下一步转向定理 4 的证明。这个证明与定理 2 的证明很类似。相对于在 A_1 上的排序$\succeq$,设 a_1 是 A_1 的第一个元素,相对于在 A_2 上的排序$\succeq$,设 p_1 是 A_2 的第一个元素。然后,在 A_1 和 A_2 上的数值函数 φ_1 和 φ_2 定义如下(对于 A_1 中的 a 和 A_2 中的 p 来说):

$$\varphi_1(a) = \begin{cases} 1 & \text{如果 } a \approx a_1 \\ -n+1 & \text{如果 } a_1 J^n a \end{cases}$$

$$\varphi_2(p) = \begin{cases} 1 & \text{如果 } p \approx p_1 \\ -n+1 & \text{如果 } p_1 J^n p \end{cases}$$

正如证明定理 2 的情况一样,这很容易表明:

$$\varphi_1(a) > \varphi_1(b), \text{当且仅当}, a > b$$

$$\varphi_2(p) > \varphi_2(q), \text{当且仅当}, p > q$$

此外,在准备证明定理 2 时所证明的预备定理 1 - 9 在当前的背景下也成立,因为它们只依赖于两个分量的二元关系。当然,对于每一个预备定理而言,严格地说,现在有一个预备定理对,一个预备定理是针对每个分量的排序。

与前面列出的预备定理 10 相对应,我们能够运用定义 9 的

公理 4,通过相同的归纳论证证明,如果 aJ^nb 和 pJ^nq,那么 $(a, q) \approx (b, p)$。第二,我们能够证明这种基本事实:如果 $(a, q) \approx (b, p)$,那么要么(i) 存在着某个 n,使得 aJ^nb 和 pJ^nq;要么(ii) 存在着某个 n,使得 bJ^na 和 qJ^np;要么(iii) $a \approx b$ 和 $p \approx q$ 。从这些结果来看,我们于是证明了下列基本结果:$\varphi_1(a) + \varphi_2(q) = \varphi_1(b) + \varphi_2(p)$,当且仅当,$(a, q) \approx (b, p)$,这就完成了定理 4 的证明。

74

3.5 部分递归函数的机器表征[①]

20 世纪初,在数学基础方面伟大的哲学成就是,把大多数数学对象表征为有某种特殊结构的集合。几十年之后,在许多方面更加深刻的成就是,证明了可计算函数表面上的不同定义都会导致相同的函数类。这里,我将形式的发展限制在来表明任何部分递归函数(*部分递归*是可计算函数的一种重要的阐述)都能通过一个无限寄存器机的程序来表征。许多其他等价的表征在文献中是已知的。

有必要承认,计算相同数学函数的两种计算机程序,可能在许多可观察的方面是不同的:计算需要的空间或时间、在程序中的行数或字符数,等等。像在本章和其他章节所考虑的各种其他表征那样,结构的同构没有必要意味着,通常也并不意味着在所有方面的观察的等价性或计算的等价性。

---

① 在不失去重要的连续性的前提下,可以略去相当技术性的这一节,除了 8.2 和 8.3 节之外。一般读者可以发现,看看头两段是有用的,因为“成为可计算的”这一概念的澄明在数学基础中是 20 世纪的伟大成就之一。

在给出有效的可计算性或部分递归性的一种形式定义之前，让我们考虑有效地可计算的函数的几个例子。[①] 我们可以从加法开始。根据后继函数 S，加法的递归定义通过下列加法函数 f 的两个方程来给出。这种递归是关于 n 的。

$$f(0,\ n) = n$$

$$f(S(n),\ m) = S(f(n,\ m))$$

例如，在用数字表示的通常符号中，

$$S(0) = 1$$

$$S(1) = 2$$

等等，这里，S 是后继函数。于是，在有限的步骤内，我们能够找到对和的一种解答，比如 $f(3,\ 2)$，或者 $3+2$。

$$\begin{aligned} f(3,\ 2) &= S(f(2,\ 2)) \\ &= SS(f(1,\ 2)) \\ &= SSS(f(0,\ 2)) \\ &= SSS(2) \\ &= SS(3) \\ &= S(4) \\ &= 5 \end{aligned}$$

在一系列的原始递归函数中，第二个函数最自然地是乘法运算 μ。 75

---

① 丘奇(Church, 1936)、克伦(Kleene, 1936)、波斯特(Post, 1936)、图灵(Turing, 1936)和马尔可夫(Markov, 1951)等人在同一个时期，以不同的方式，首先描述了部分递归函数的类。

$$\mu(n, 0) = 0$$

$$\mu(m, S(n)) = f(m, \mu(m, n))$$

在这里,没有从形式上定义这些原始递归函数,因为形式定义是相当复杂的,但是,在递归是基本的地方,它们是部分递归函数的最重要的子集,是总函数。原始递归函数的几个其他例子如下。第一个是求幂。

$$p(n, 0) = 1$$

$$p(m, S(n)) = \mu(p(m, n), m)$$

阶乘函数 $n!$ 恰好是:

$$0! = 1$$

$$(n+1)! = (n+1)n!$$

$n \dot{-} 1$ 指算术的减去 1。在递归意义上定义它的根据是:

$$0 \dot{-} 1 = 0$$

$$S(n) \dot{-} 1 = n$$

而且,一般情况下,

$$m \dot{-} 0 = m$$

$$m \dot{-} S(n) = (m \dot{-} n) \dot{-} 1$$

归因于阿克曼(Ackermann)的迭代或一般的指数函数,提供了不是原始递归而是部分递归函数的一个例子。

$$f(0, 0, n) = n$$

$$f(0, m+1, n) = f(0, m, n) + 1$$

$$f(1, 0, n) = 0$$

$$f(p+2, 0, n) = 1$$

$$f(p+1, m+1, n) = f(p, f(p+1, m, n), n)$$

注意，

$$f(0, m, n) = n + m$$

$$f(1, m, n) = n \cdot m$$

$$f(2, m, n) = n^m$$

不是原始递归的而是部分递归的第二个函数是通过下列三个方程下定义的： 76

$$U(0, n) = n + 1$$

$$U(m+1, 0) = U(m, 1)$$

$$U(m+1, n+1) = U(m, U(m+1, n))$$

但是，我们将不提供证明这个事实的方法。

定义 1. 其自变量和值是自然数的一个函数，是部分递归的，当且仅当，它能够通过应用Ⅳ、Ⅴ和Ⅵ的一个有限数，根据函数Ⅰ、Ⅱ和Ⅲ来获得：

Ⅰ. $S(x_1) = x_1 + 1$（后继函数）

Ⅱ. $0^n(x_1, \cdots, x_n) = 0$（常数函数）

Ⅲ. $U_i^n(x_1, \cdots, x_n) = x_i$（投影函数）

Ⅳ. 如果 $h, g_1, \cdots, g_m$ 是部分递归的，那么函数 f 也是如此，因此，

$f(x_1, \cdots, x_n) = h(g_1(x_1, \cdots, x_n), \cdots, g_m(x_1, \cdots, x_n))$（合成函数）。

Ⅴ. 如果 g, h 是部分递归的，那么函数 f 也是如此，因此，

$f(0, x_2, \cdots, x_n) = g(x_2, \cdots, x_n)$

$f(z+1, x_2, \cdots, x_n) = h(z, f(z, x_2, \cdots, x_n), x_2, \cdots, x_n)$（原始递归的）；

Ⅵ. 设 μy 意味着“最小的 y,使得”。如果 g 是部分递归的,那么函数 f 也是如此,因此,

$$f(x_1, \cdots, x_n) = \mu y[g(x_1, \cdots, x_n, y) = 0]$$

这里,如果没有这样的 y 存在,那么对给定的论证来说,没有定义 f。

现在,自然数的可计算函数概念等同于部分递归函数概念。由于前面提到的丘奇、克伦、马尔可夫、波斯特和图灵等人的工作,许多等价性是已知的。

我们首先运用由谢泼德森(Shepherdson)和斯特吉斯(Sturgis)提出的更容易理解和更直观的无限寄存器机(URMs),而不是运用部分递归函数直接表征的通用图灵机(UTMs)。我追溯它们的发展。

无限寄存器机(URM)。我首先从某些符号开始: <n>是寄存器 n 在执行一条指令前的内容;<n'>是执行一条指令后的内容。

定义 2. 一个 URM 有,(i) 编号为 1、2、3…的寄存器的一个可数序列,每个寄存器都能存储任何一个自然数,但每个程序只用一个有限数的寄存器;(ii) 六个基本指令:

(a) $P(n)$: 加 1, $<n'> = <n> + 1$;

(b) $D(n)$: 减 1, $<n'> = <n> - 1$ if $<n> \neq 0$;

(c) $0(n)$: 清空寄存器, $<n'> = 0$, 即只有数 0 被存储在这个寄存器中;

(d) $C(m, n)$: 从寄存器 m 拷贝到 n, $<n'> = <m>$;

(e) $J[E_1]$: 跳转到退出 1;

(f) $J(m)[E_1]$：如果 $<m>=0$，跳转到退出 1。[①]

77

当我们用(e)和(f)的程序时，形式 $J[3]$：表示跳转到第(3)行。如果跳转到一个不存在的行，机器就停止了。因此，E_1 是跳转到的行数的一个变量。

作为一个例子，下列程序把寄存器 1 中存储的“一半的”数放到寄存器 2 中，其余的数在寄存器 3 中。

1. $J(1)[8]$　如果 $<1>\neq 0$，那么到达第 2 行。在其他情况下，到达第 8 行(即，停止)

2. $D(1)$　寄存器 1 的数减去 1

3. $P(2)$　寄存器 2 的数加上 1

4. $J(1)[8]$

5. $D(1)$

6. $P(3)$

7. $J[1]$　跳转到行 1

我们现在定义 URM 的行和程序。

定义 3. (一个 URM 的一个程序的)一行，要么是由 $n\geqslant 1$ (行数)的自然数和指令(a)～(d)之一组成的一个有序数偶，要么是由 $n\geqslant 1$ 的自然数、指令(e)或(f)之一和一个自然数 $m\geqslant 1$ 组成的有序三元组。

定义 4. (一个 URM 的)一个程序是 l 行的一个有穷序列，使得

(i) 第 i 行的第一项是 i；

(ii) 数 m 是行的第三项，它满足条件：$1\leqslant m\leqslant l+1$。

我们也把程序(programs)看成是例行程序(routines)。一

---

① 注意，有一个“空的寄存器”是方便的，这意味着只有 0 储存在它当中，因为如果根本没有储存数，那么一定要为 $<n>-1$ 和其他表达提供特殊定义。

个 URM 如何遵守一个程序在直观上是显而易见的，而且将不从形式上下定义。

子程序(subroutines)像程序那样下定义，除了

(i) 子程序可能有几个退出；

(ii) 三元组的第三项可以在范围 E_1，…，E_k 之内，同时在给定程序中为这些变量赋值。

子程序$\bar{0}(n)[E_1]$的事例：清空寄存器 n 并进入退出 E_1。

1. $J(n)[E_1]$

2. $D(n)$

3. $J[1]$

很容易证明子程序的可分配性。

定义 5. 设 f 是 n 个自变量的一个数论函数。于是，f 是通过一个 URM 可计算的，当且仅当，对于自然数的每一个集合 $\{x_1, \cdots, x_n, y, N\}$（其中，$y \neq x_i$，$i = 1, \cdots, n$，而且 $x_1, \cdots, x_n, y \leqslant N$）来说，都存在着一个例行程序 $R_N(y = f(x_1, \cdots, x_n))$，使得如果$<x_1>$，…，$<x_n>$是寄存器 x_1，…，x_n 最初的内容，那么

(i) 如果 $f(<x_1>, \cdots, <x_n>)$ 是未定义的，那么这个机器将不会停止；

78 (ii) 如果 $f(<x_1>, \cdots, <x_n>)$ 是已定义的，那么机器将停止在$<y>$，寄存器 y 的最终内容等于 $f(<x_1>, \cdots, <x_n>)$，并且与开始时一样，含有除 y 之外的所有寄存器的最终内容 1，2，…，N。

因为对于用来定义部分递归函数集的六个表达式(Schemata)中的每个表达式来说，主要定理的证明是由编写子程序构成的，所以，对子程序作出更多的评论可能是有用的。

从逻辑的观点来看，子程序类似于在一个理论中的定义。

只根据成员关系的原始概念,集合论的数学发展,不仅是难以容忍的,而且在实践中是不可能的。这种评论同样适用于编程。用单纯的计算机语言编写复杂程序是难以容忍的。从形式的观点来看,定义理论得到了全面的发展,而且很容易给出定义规则,这些规则保证,在所引入的定义中,没有由特性或不受怀疑的特性所引发的问题。好的定义应该满足无创造性和可排除性的标准。无创造性标准要求,只因为引入这个定义,就不可能在原始符号中证明任何一个所阐述的定理。这里用一个简单的例子来阐明所需要的这个标准。我们只把结合的公理看作我们理论的惟一公理:

(1) $$x \circ (y \circ z) = (x \circ y) \circ z,$$

我们现在引入常数 e 的"定义":

(2) $$x \circ e = x$$

从这个定义来看,但不是从上面的公理来看,我们能证明:

$$(\exists y)(\forall x)(x \circ y = x)$$

因而所提供的定义是不能令人满意的,因为它违反了非创造性标准。从直观上看,显然,(2)作为与(1)相结合的定义是不令人满意的,因为它把概念内容增加到这个理论当中。

同样的方式,一个子程序一定也是非创造的。通过运用这个子程序的一个程序是可计算的任何一个函数,不用这个子程序的一个程序一定也是可计算的。

可排除性标准要求,在所有的语境中,一个定义的符号是根据前面的定义或原始符号可排除的,当然还要求,在用了这个符号的陈述和除去这个符号相对应的陈述之间,等价性成立。[关于这两个定义标准的更详尽讨论的一个参考文献是苏佩斯的著作(1957/1999,第 8 章。)]

下面是不满足可排除性标准的一个子程序对的一个例子。

$0_1(n)$:	$0_2(n)$:
1. $0_2(n)$	1. $0_1(n)$
2. $P(n)$	2. $D(n)$

换言之,根据 $0_1(n)$定义子程序 $0_2(n)$,根据 $0_2(n)$定义 $0_1(n)$,这样,我们陷入了一个恶性循环:不可能根据这六个基本指令来定义或编写 $0_1(n)$或 $0_2(n)$的程序。根据下列要求,恰好像循环定义那样,禁止循环子程序:即在固定的线性次序中引入子程序,并且第 n 个子程序在其程序中只能用基本指令和前边的 $n-1$ 个子程序。(显然,这种线性要求可能被弱化。)

79

我们现在转向表征定理。

定理 1. (机器表征)一个数论函数是部分递归的,当且仅当,它是通过一个 URM 可计算的。

证明: 证明通过一个 URM 可计算的函数是部分递归的是不困难的:检查一下,在由部分递归函数给出的寄存器中,一个程序的每一行都发生了一种变化,然后表明,由于部分递归函数集的闭合性,所以变化的序列是部分递归的。

由于一个 URM 指令是很简单的,所以另一种方式的论证是令人感兴趣的部分。这点应归功于谢泼德森和斯特吉斯(1963)的证明,比图灵机的相应证明更简单。对于定义部分递归函数的定义 1 的表达式Ⅰ—Ⅵ来说,它恰好是由编写 URM 的子程序组成的。

Ⅰ. 子程序,$R_N(y = S(x))$

1. $C(x, y)$

2. $P(y)$

Ⅱ. 子程序,$R_N(y = 0^n(x_1, \cdots, x_n))$

1. $0(y)$

Ⅲ. 子程序，$R_N(y = U_i^n(x_1, \cdots, x_n))$

1. $C(x_i, y)$

Ⅳ. 合成函数的运算，$R_N(y = f(x_1, \cdots, x_n))$，这里，我们用子程序 g_i、h 定义：

f：$f(x_1, \cdots, x_n) = h(g_1(x_1, \cdots, x_m), \cdots, g_m(x_1, \cdots, x_n))$

1. $R_{N+1}(N+1 = g_1(x_1, \cdots, x_n))$

⋮

m. $R_{N+m}(N+m = g_m(x_1, \cdots, x_n))$

$m+1$. $R_{N+m}(y = h(N+1, \cdots, N+m))$

Ⅵ. 子程序，$R_N(y = f(x_1, \cdots, x_n))$，这里，我们通过Ⅵ用子程序 g 定义 f：$f(x_1, \cdots, x_n) = \mu y[g(x_1, \cdots, x_n, y) = 0]$

1. $0(y)$
2. $R_{N+1}(N+1 = g(x_1, \cdots, x_n, y))$
3. $J(N+1)[6]$
4. $P(y)$
5. $J[2]$

注意，在指令 1 之后，$\langle y \rangle = 0$。如果 $\langle N+1 \rangle = 0$（第三步），那么我们结束，并进入了不存在的第 6 行。如果不是，我们使 $\langle y \rangle$ 加 1，然后，回到第 2 行，对最小的 y 再试一次。

表达式Ⅴ稍微更加复杂一点。有两个符号问题需要处理。第一，我们现在只数出必要的行数——子程序或跳转。第二，我们用一个符号代表指令的重复。如果 I 是一个退出指令或不影响寄存器 n 的子程序，那么 $I^{\langle n \rangle}$ 代表 $\langle n \rangle$ 次执行 I 的结果和使 $\langle n \rangle$，即寄存器 n 中的数还原为零的结果。 80

Ⅴ. 子程序，$R_N(y = f(x_1, \cdots, x_n))$，这里，我们根据Ⅴ用子程序 g 和 h 定义 f：

$$f(0, x_2, \cdots, x_n) = g(x_2, \cdots, x_n),$$

$$f(z+1, x_2, \cdots, x_n) = h(z, f(z, x_2, \cdots, x_n), x_2, \cdots, x_n):$$

1. $R_N(y = g(x_2, \cdots, x_n)), 0(N+1)$

2. $\{R_{N+2}(N+2 = h(N+1, y, x_2, \cdots, x_n)), C(N+2, y), P(N+1)\}^{<x_1>}$

3. $C(N+1, x_1)$

任意有穷字母表上的部分递归函数。为了进行更一般的设置和拥有不用“抽象”数作为自变量的一种阐述，我们可以在一个有穷字母表 V 上定义部分递归函数。V 是一个有穷集合，V^* 是 V 元素的所有有穷序列的集合；在当前的语境中，我们把 V^* 的元素称为词(words)。下面，变量 a_i 的值域在 V 上，变量 $x, y, x_1, \cdots, x_n$ 等的值域在 V^* 上。同样，xa_i 是 x 和 a_i 的并置。

表达式 Ⅰ′—Ⅵ′如下：

Ⅰ′. $S_{ai}(x) = xa_i$

Ⅱ′. $0^n(x_1, \cdots, x_n) = \varnothing$

Ⅲ′. $U_i^n(x_i, \cdots, x_n) = x_i$

Ⅳ′. ＝Ⅳ |（合成函数的运算）

$$f(x_1, \cdots, x_n) = h(g_1(x_1, \cdots, x_n), \cdots, g_m(x_1, \cdots, x_n))$$

Ⅴ′. $f(\varnothing, x_2, \cdots, x_n) = g(x_2, \cdots, x_n)$

$$f(za_i, x_2, \cdots, x_n) = h_i(z, f(z, x_2, \cdots, x_n), x_2, \cdots, x_n)$$

$$i = 1, \cdots, s$$

Ⅵ′. $f(x_1, \cdots, x_n) = \mu_i y[g(x_1, \cdots, x_n, y) = \varnothing]$

这里，$\mu_i y$ 意指只由 a_i 组成的最短的词。

我们与以前一样的方式在 V 上定义 URM。

基本指令能够被减少到三个，在前面的情形中也是这样的。

a′) $P_N^{(i)}(n)$：把 a_i 置于 $\langle n\rangle$ 的最右边；

b′) $D_N(n)$：如果 $\langle n\rangle\neq\varnothing$，删除 $\langle n\rangle$ 的最左边的字母；

c′) $J_N^{(i)}(n)[E_1]$：如果 $\langle n\rangle$ 从 a_i 开始，跳转到退出 1。

(如前所述，参数 N 指除了上面的运算之外，其余没有变化的寄存器。)对于数值的部分函数来说与前面一样的表征定理成立。

下面，我们转向只用一个寄存器(SUM)的图灵机(TMs)，而且在被处理的 V^* 的词中存储信息或数据。① 省略了对还原的适当性的证明。就标点符号而言，我们把标点加到 V 中。一个 SRM 在 $V\cup\{,\}$ 上的三个指令是：

81

a″) $P^{(i)}$：$x\rightarrow xa_i$(把 a_i 加到 x 最右边)；

b″) D：$a_ix\rightarrow x$(删除第一个字母)；

c″) $J^{(i)}[E_1]$：如果 x 从 a_i 开始，跳转到退出 1。

设 $V=\{1\}$ 并把'看成是 0。于是，一个 SRM 的结果表明，自然数的任何部分递归函数通过下列机器都是可计算的：这个机器具有单向带(在它上面打印 x)、两个带符 0 与 1 以及两个头，左边的读头和右边的写头。

寄存器机的许多其他细节是在谢泼德森和斯特吉斯的著作中提供的(1963)。关于图灵机的可计算的详细结果的资料很多，例如，戴维斯(Davis，1958)、罗杰斯(Rogers，1967)的文献，在更高的层次上，还有索阿雷(Soare，1987)的文献。

3.6 心理表征的哲学观点

在哲学中，心理表征概念的发展历史悠久而错综复杂，如果

---

① 对图灵机的更详细的处理在第 8 章开始定义 8.2.9 时提供。在 8.3 节中，寄存器机还有另外的用途。

我们不打算对其进行详尽而准确的叙述，那么可能有所启发和相关的是，评论从亚里士多德以来的某些问题和标准概念。主要的问题不是评价已有的分析是否正确或批判其是否适当，而是反思在这一背景中是否有同构概念，正如在相像(likeness)或类似性(resemblance)这些概念上所反映出的那样。我也会对这些概念如何与表征联系在一起作出评论。

亚里士多德。概括地说，定义感知的特征就是接受物体的形式，这是亚里士多德最著名的知觉观。就他所提供的细节而言，这根本不是一种坏的观点，即使从现代的观点来看，这很显然不可能是完全正确的。

亚里士多德讨论知觉的背景和因此隐含了讨论心理表征的背景是，他把形式(form)与质料(matter)区分开来，这种区分适用于所有类型的对象和现象。例如，一把斧头的形式不同于接受这种形式的质料。这种区分在下列意义上也是相关的：例如，一间房屋的质料可能是由砖块构成的，依次，这些砖块有它们自己的形式和构成它们的质料。我在苏佩斯的文章(1974c)中对这个一般的质料概念的持续的生命力作出了辩护。

亚里士多德在他的知觉理论中利用了形式与质料的区分。正如下面引自第二本书《论灵魂》(*De Anima*)中的一段所描述的那样，感觉器官的作用是潜在地得到一个物体的形式，而不是它的质料。

> 我们必须把通常情况下的每一种感觉都理解为是真的：(1) 这种感觉是，在没有质料的条件下，善于接受可感知对象的形式，正如蜡制的印章是接受了没有铁或金的环状物的印象和接受了金或青铜的印象，而不是接受金或青铜本身；因此，在每一种情况下，感觉都会受到具有颜色、或气味、或声音的东西的影响，但通过感觉不是拥有一种特殊

的同一性本身，而是拥有某种性质本身，并借助于它的方法；(2) 感觉器官的原义是其中隐含了这种潜在性。

(De Anima, 424a17 - 25)

因此，如果我们在感知蜡烛时，得到的正好是蜡烛的形式，而不是它的质料，那么我们就有了实物蜡烛的一个表征，在心灵里，这种表征与实物的蜡烛是同构的，恰好是由于形式的相同性(sameness)。这里相关的观点是，形式的相同性概念与我一直作为表征的中心概念来论证的同构概念非常符合。形式概念在下列意义上很好地捕获到了同构概念：形式因此不是涉及蜡烛的所有特性，而是只涉及被视觉、触觉等各种不同感官感知到的特性。正如在前面的讨论中已经注意到的那样，对所考虑的特性的这种限制是任何一个可使用的同构概念的特征。 82

读了上面的一段引文和我的评论之后，有人可能会细想，这并不是很丰富的心理感知理论，只是直接感知理论。然而，亚里士多德继续把关于形式的相同思想延伸到智力。这里不可能详细地讨论他提出的理论中包括的所有微妙的思想，但是下面的一段引文明确地表明了，思考与判断的灵魂部分是如何以与形式相同的方式起作用的，因而，特有的同构概念还是可应用的：

论及灵魂的那个部分(不管它在广延的空间中是可分离的，还是只在思想中是可分离的)，根据这个部分，灵魂进行认识和思考，我们不得不考虑，它的显著特征是什么，如何产生思考。如果它类似于知觉，那么它一定要么是一个过程，在这个过程中，灵魂受到了能思考的东西的影响，要么是同样类型的别的东西。于是，这部分必须(尽管是被动的)善于接受一个对象的形式，也就是说，在潜在的意义上，必须与它的对象一样，尽管并不是与它同一：像敏感的是

> 可感觉的一样，心灵也一定是可以思考的。
>
> ……因此，心灵也不可能拥有除了它的接受能力之外的特征。于是，灵魂的那个部分，我们称之为心灵（我所说的心灵的意思是，灵魂据其进行思考和形成判断的那个部分），只有当它进行思考时，它才是实际存在的。这样，没有理由假设它与身体混合在一起；因为在那种情况下，它在某种程度上成为定性的，比如，热或冷，或者，它甚至会有某种器官，像敏感力有器官一样，但事实上，它却没有。有句话说得好，灵魂是形式的处所，除了这不适用于作为整体的灵魂但只应用于它的思考能力方面之外，并且这些形式不是现实地而只是潜在地占据了灵魂。
>
> （De Anima，429a10－18，a23－30）

形式是智力沉思的焦点，但不能离开个人的身体而存在，这是亚里士多德思想的一个明确的组成部分。根本没有单独的柏拉图式的形式的宇宙。对亚里士多德的这种观点的更加明确的陈述是圣·托马斯·阿奎那在下面一段引文中作出的。

> 我这么说，是因为有些人坚持认为，形式只属于种，而质料是个体的一部分，不是种的一部分。情况并非如此，因为种的本性属于定义的所指，而且，在自然物中，定义不是只表示形式，而是表示形式和质料。因此，在自然物中，质料是种的一部分；确实不是具体质料（signate matter），这是个体化原理，而是表示共同的质料（common matter）。因为正如由这个灵魂、这个肉体和这些骨头构成的属于这个特殊的人的本性一样，由灵魂、肉体和骨头构成的也属于人的本性；因为共同属于包含在某一给定物种中的所有个体的本质的无论什么东西，一定也属于这一物种的本质，一

定也属于这个种的本质。

(*Summa Theologica*, I, Q. 75. Art. 4)①

笛卡儿。笛卡儿并没有坚持由亚里士多德提出经过阿奎那进一步发展了的所有的心理学方法,他仍然坚持下列学院观点:一个对象的视知觉反映了这个对象的形式。这来自这样的物理事实:"从它的身上反射出来的光,描绘了它的两个图像,在我们的每一只眼里都有一个图像。"因此,他在这里用生理光学的自然事实最后说明了在灵魂中的图像或形式。下面是我知道的《论灵魂的激情》(*Passions of the Soul*)一书中的一段。 83

> 因此,例如,如果我们看见某个动物向我们走近,从动物身上反射出来的光描绘了它的两个图像,在我们的每一只眼里都有一个图像,而且,这两个图像通过视神经在大脑内侧面向空腔的表面形成了两个其他图像;然后,从此,借助于充满大脑空腔的活力,这些图像如此辐射到由这些活力所包围的那个小密封体内,以致形成一个图像的每个点的运动,都趋于这个密封体的相同的点,向着这个密封体的运动形成了另一个图像的点,这代表了这个动物的相同部分。用这种方法,在大脑里的两个图像只在这个密封体内形成一个图像,这个图像立即作用于灵魂,促成了看到这个动物的形式……
>
> (Article XXXV)

一般情况下,笛卡儿当然拒绝接受许多学术教条,但在他的视觉心理学中,完全停留于灵魂的自然思想或拥有图像的心灵,据此这些图像只反映了所看见的对象的形式。这是亚里士多德知觉

① 译自《圣·托马斯·阿奎那的主要著作》(1944)。

学说的核心,在笛卡儿的思想中,尽管有许多其他变化,但是这个核心仍然没有变。根本的变化是由休谟带来的。

休谟。由于种种原因,休谟在《人性论》(*Treatise of Human Nature*)(1739/1951)中极力反对在外部世界中的对象和内心意象之间的任何直接的心理表征概念。他对这些问题有许多论述,我只引用和评论关键的几段。

要强调的第一点是,对于休谟来说,心灵的机制是联想,而且他骄傲地并且在许多方面正确地认为,联想在心灵理论中所起的作用,与引力在牛顿的世界体系中所起的作用是一样的。换言之,在落体落向地面的问题和行星在宇宙中运动的问题中,因果性原理是两个物体之间的万有引力,除此之外,就不可能找到说明。在心灵里,这种机制是联想的机制,两者都在印象之间,它们在许多情况下可能与感知和思想相一致,它们总是从印象中得到的。印象和思想有三种特征会引起联想。它们是类似性、时间或位置的邻近(contiguity)以及因果关系。对于休谟来说,重要的是强调产生印象的联想的惟一特征是类似性。所以,所有的推理,例如,关于因果关系的推理,一定都在思想的联想之间。指定了这三种特性之后,休谟定义了类似性:

> 这些性质有三种,即类似、时空接近以及因果关系……。联想是从这些性质中产生出来的,而且心灵通过这些性质追求把一种思想传递给另一种思想的这种方式。
>
> 84 ……我认为,证明这些性质在思想中和当一种思想的表象自然会产生另一种思想的表象时,产生了一种联想,完全是没有必要的。很显然,在我们的思维过程中和我们思想的经常的转变中,我们的想象容易从一种思想变到与它类似的任何另一种思想,并且只有这种性质才是这种想象

> 力的充分的联系和联结的原则。
>
> (*Treatise*: p. 11)

这一段的其他部分对邻近作出了相似的辩护,并表示,后面会更加详尽地讨论原因与结果。

与这里的讨论相关的重要观点是,休谟把类似性推到显著的位置,因为正如他所指出的那样,只有它是印象之间的一种性质。当然,类似性正好是休谟接近于作为表征基础的现代的同构概念的概念。

我前面已经说过,休谟否认在不连续的知觉和被认为产生了那些知觉的真实对象之间有直接的类似性,但他确实辩护了实物的实在概念和它们随时间的连续同一性。他关于这些问题的冗长而复杂的论证是他的《人性论》第一卷中最令人感兴趣的部分。我想指出他运用类似性概念或同构概念的各种不同方式。下面的一段特别引人注目,因为他通过运用下列事实,论证了我们对有点像太阳或其他很稳定的对象的知觉的恒常性,这种事实是:我们的连续知觉,尽管有中断,但彼此类似,因而,正如他详细地论证的那样,证明了关于对象的跨时间的同一性的推理是正当的。下面是关于类似性的这种用法的关键一段。

> 当我们在特定的印象中一直习惯于观察到一种恒常性时,当我们发现,比如,对太阳或海洋的知觉,缺乏或消除了相像的部分和相像的秩序(像初始表象一样)后,返回给我们时,我们不倾向于把这些中断的知觉看成是不同的,(它们实际上如此)而是相反,由于它们的类似性,把它们看成在个体意义上是相同的。但是,当它们存在的这种中断与它们的完全同一相反,并使我们把初始印象看成是已消除的,把第二印象看成是新创造的时,我们发现,我们自己有

> 点不知所措，陷入了一种矛盾之中。为了使我们摆脱这种困境，我们尽可能地掩饰这种中断，或更确切地说，通过下列假设完全排除这种困难：这些中断的知觉通过我们无法觉察到的一种实际存在联系起来。这种连续存在的猜测或思想，从这些间断印象的记忆中和从它们提供给我们的假设它们是相同的倾向性中，获得了效力和生命力；并且，按照前面的推理，信念的真实本质存在于概念的效力和生命力之中。
>
> （*Treatise*：p. 199）

休谟并没有就此停止。他用了好几页，在从中断的知觉到对象随时间的同一性的推理中，说明了这种设置的四个不同方面。这是在《人性论》中最持久不变的论证之一，也是一种奇妙的哲学分析。他首先说明了对象的个体化原理的基础，然后，他在这里提出了完全接受同构和相关的不变性概念的某种观点。下面是讨论这个原理的最短的一段。

85

> 因此，在不放弃这种观点，也没有责任形成多样的或大量的思想之前提下，个体化原理只不过是心灵通过假设的时间变化能在其不同时期的存在中追踪的任何一个对象的不变性和不间断性。
>
> （*Treatise*：p. 201）

这个不变性原理是随时间不变的一个原理，证明了对象的恒常性的推理是正当的。第二个说明是，从恒常性推出数值完全同一的概念。下面是这种冗长论证的最后一段。休谟很好地总结了我们为什么开始相信对象是连续存在的。

> 我们的记忆向我们呈现了在不同的时间间隔和经过很多中断之后，获得的彼此完全类似的知觉的大量事例。这

种类似性向我们提供了一种倾向，把这些中断的知觉看成是相同的；还有一种倾向是，通过一种连续的存在性把它们联系起来，以便证明这种同一性是正当的，以及避免出现这种矛盾：这些中断的知觉表象似乎必定把我们包括在内。于是，这里我们有一种倾向，虚构所有感知体都是连续存在的；而且，因为这种倾向是从这种记忆的某些活生生的印象中产生的，这就对那种幻想赋予一种生命力；或换句话说，使我们相信物体是连续存在的。如果我们有时把一种连续存在归因于对我们来说是全新的对象，而且，我们没有经历过它们的恒常性和一致性，那么，这是因为，它们本身呈现给我们感官的方式，类似于稳定的和一致的对象的方式；这种类似性是推理和类比的一种来源，并致使我们把相同的性质归于相似的对象。

(*Treatise*: pp. 208 - 209)

在这一段的结尾，类似性作为休谟的同构概念所起的关键作用，自始至终都是明显的。

对于休谟来说，在这一点上，这种论证的结论不会就此为止。他对外在对象的存在性问题的详尽而复杂的攻击，包括后面阶段的人格同一的问题在内，延续了另外的六十页一直到《人性论》第一卷的结束。探索这些细节是在转移注意力，但是，有一两个观点与他在类似性概念方面所用的同构的考虑直接相关。在第一卷的其余部分，类似性、邻近以及原因与结果作为思想或印象之间的重要关系保持了它们的中心地位。这种引申用法的一种好例子是，根据类似性对记忆的描述。

这种记忆为什么只是一种能力(我们通过这种能力形成过去的知觉意象)呢？当一种意象必定类似于它的对象

> 时，频繁地把这些类似的知觉置于思想链中，一定不会更容易把这种想象力从一种联系传递给另一种联系，也不会使这个整体看起来像一个对象的连续吗？那么，在这种特殊情况下，记忆不仅发现了同一性，而且通过在这种知觉中产生的类似关系，促进了它的产生。
>
> (*Treatise*: pp. 260 - 261)

在这方面，休谟的观点，正如他如此经常评论的那样，类似于这种通俗的见解：即，记忆的重要特征恰好是产生类似于过去知觉图像的特征。在我自己的观点中，很难有与此完全不同的一般的记忆观。当然，细节是极其复杂的，而且现在也已经有许多研究。这是对休谟的某些基本机制的描述，这些机制既促进了也曲解了准确的记忆。我想到了这方面的广泛研究：联想有助于准确地回忆，也促进了误导性的干涉和不准确的方式。令人惊奇的是，在用现代术语描述特有的神经机制的问题时，这种记忆理论的进展是何等地微不足道。一个好的例子是，根据从记忆中挽回的那些项，对这种机制进行详细的物理学或化学描述——例如，在阅读或听讲时的词语。在记忆的所有机制中，这是在每一种情境中那些最不变的要求之一，重要的是不要对这一点产生误解。有关描述记忆的研究范围很广泛。最近的一个出色评论可参见图尔文和克雷克的著作(Tulving and Craik, 2000)。仍然缺乏的是，更深刻理解支持记忆认知活动的物理学与化学机制。(第 8 章最后一节关于语言的脑电图表征讨论了这些问题。)

86

为了结束休谟的复杂的进一步论证，我只提到一般的思路，在讨论表征时乃至更多地讨论同构概念时，这种思路是很重要的。休谟对外在的同一性或人格同一性的十分简单的思想的首要攻击，是由指出我们如何在日常语言和经验中使用相同性概

念构成的。例如,我们会说,一艘船是相同的,即使修理或置换了许多零件。我们承认,植物、动物和小孩子在短期内会发生很大的变化,可是,我们把他们看成是同一棵树、同一匹马和同一个孩子。在某些情况下,也可能有相当大的变化。他提到了以不同的风格彻底地重建的一所教堂。我们仍然可以把它看成是同一教区的教堂进行谈论。这种观点是,相同性概念(我们认为它是同一性的替代品)实际上与知觉、用法、意向和习惯的许多方面相关。它不同于类似性。对于休谟来说,类似性是更基本的,而且,在任何一种可靠的哲学中,它依然如此。真正的观点是对同一性的主张的怀疑,并承认,关键之处是,在不同的时间和地点,经验之间的近似的同构或类似性概念作为我们建立对个人、地点和事物的更精致的理论的基础。

康德。康德是打破同构与表征之间长久联系的第一位主要的哲学家。他对表征有许多论述,但通常没有提供关于同构概念的看法。令人注目的是,在从亚里士多德到休谟关于这些问题的悠久的思想史中,还没有人更多地注意到,康德实际上是从作为核心的可行的同构概念转向他的表征概念的第一位主要人物。由康德摧毁的这种联系的生命力,在后来的哲学中并没有得到彻底地恢复,后来的哲学实际上充满了对亚里士多德建立在可感觉到的和可理解的形式概念之基础上的知觉理论的怀疑,对休谟著作中达到顶点的经验传统,同样也是如此。

我对历史的解读是,休谟和康德两人都把牛顿的自然哲学及其所取得的成就看成是他们的科学典范,正如在牛顿的《自然哲学的数学原理》中如此详细地阐述的那样。但是,他们抓住了完全不同的指向。休谟的目标是提出人类本性的经验理论,这种理论作为一门科学能与牛顿理论相媲美,休谟的联想概念在这个理论中所起的作用相当于引力在牛顿体系中所起的作用。 87

对于休谟和牛顿来说，当前根本没有一种说明有可能超越联想对人类本性的说明，或者万有引力(等力)对自然界的说明。在休谟的术语中，联想和万有引力两者都是原始性质。康德拥有一个不同的议程。他关注的是，休谟的科学进路只通向偶然的因果关系，更一般地说，只通向世界的偶然知识。对他而言，牛顿工作的优点是清楚而明确地确立了必然的自然规律的形式。为了在它的最佳意义上提供知识，这样一种必然基础是康德追求的目标。在阐述他的精致体系的细节时，他更多地论及表征，很少论及同构。休谟希望为他认为的心灵的运行过程提供一种真正的经验说明。相反，康德则关心通过先验的论证来确定，心灵如何能够确立关于世界的必然真理，而且这些真理如何一定是有限的，如何必须与形而上学的思辨隔离开来。

在无疑是意指休谟时，康德在《纯粹理性批判》(1781/1997)的标题为"先验分析"部分是这么说的：

> 确实，它只是一个经验的规律，根据这个规律，彼此跟随或伴随的表征，最终彼此关联，因此而处于一种联结中，根据这种联结，即使没有对象，这些表征之一也根据不变规则把心灵带向另一个表征。然而，这个再生规律预设了，表象本身实际上受到这样一个规则的支配，而且在它们的多种形式的表征中，产生了遵从特定规则的伴随或跟随；因为如果没有这个假定，我们的经验想象力就决不会开始做出符合其能力的任何事情，因而像死的一样仍然隐藏在内心，并且对我们来说是未知的能力。
>
> (*Critique*：A100 - A101)①

① 在最近古耶尔(Guyer)和伍德(Wood)的翻译中是第229页。

两页之后，康德直接表达了他自己对休谟完全运用相继知觉之间的类似性允许作出“永久对象的幻想”的反对意见。他给出的一种先验的论证是，一定存在某些东西，在内心里，这些东西是更加先天地给定的，在内心里超越了只是类似知觉的连续。(A102)[①]

康德在阐述他的体系时，存在着一个我不希望忽略或低估的重要而有意义的问题。在他的《自然科学的形而上学基础》(*Metaphysical Foundations of Natural Science*)(1786/1970)一书中，他确实以一种极好的方式试图为牛顿的科学提供先天综合的基础。他主张以先天综合的方式导出“动力学的形而上学基础”这一章的命题Ⅰ，这个命题断言，物质不是根据它的纯粹的存在，而是根据一种特殊的运动力，充满了空间，然后，在对这个命题的说明中，他引入了吸引力与排斥力的两种可能性。 88
他不试图断言，他能够在这个先天综合框架内导出实际的万有引力的反平方定律。

最后对康德的观点作出评论。我用《纯粹理性批判》中的“概念分析”一开始有启发的引文来结束这些评论。它表明，康德是如何不关注用联想来理解怎样在经验上建构概念。用康德的技术性的短语来说，正是统觉的统一性把一切拉在一起，并允许把一个概念与另一个概念联系起来，但是这种统觉的统一性不是心灵运行方式的一种经验特征。它是必须根据知性[②]或知

---

① 在最近古耶尔(Guyer)和伍德(Wood)的翻译中是第230页。

② 这里的英语词是“understanding”，这个词在科学哲学中一般译为“理解”，本文基本上采纳了这种译法，但是，在关于康德观点的语境中，英语的“understanding”对应于德语的“Verstand”，康德的《纯粹理性批判》是从德文译过来的，通常把“Verstand”译为“知性”。因此，为了保持与国内的习惯译法一致，本文只在有关康德的引文与评论中译为“知性”。——译者

识的要求所提出的一个先天概念。注意在这段引文的最后一句,我们获得了对表征和判断的表征,但是,在这些或那些表征之间的关系中,根本没有关于任何一种同构的详细断言。

> 上面只是消极地把知性解释为是一种非感性的认识能力。既然我们不可能独立于感性来享有直观。因此,知性就不是直观的能力。但除了直观之外,再没有别的认识类型,除非通过概念。因此,每一个知性的认知,起码是人类知性的认知,都是通过概念的一种认知,不是直觉性的,而是推论性的。一切直观,作为感性的东西,都取决于刺激(affection),概念则取决于功能。然而,我把功能理解为,在一个共同表征下整理不同表征的行动的统一性。所以,概念以思维的自发性为基础,如同感性直观是以印象的接受性为基础一样。现在,知性只能用这些概念作出判断,而不可能对它们作别的运用。既然除了仅有的直观外,没有一个表征与对象直接相关,因而,一个概念决不会与一个对象直接相关,而总是与对象的某一个其他表征相关(无论这种表征是一种直观,还是本身已经是一个概念)。为此,判断是对一个对象的间接认识,从而是对对象的一种表征的表征。①
>
> (*Critique*: A68 - B93)②

詹姆斯。在康德之后的一个世纪和休谟之后的一个半世纪的作品中,詹姆斯更多地谈到了联想律的本性。我这里只限于他的《心理学原理》(*Principles of Psychology*)(1890/1950)的

---

① 这一段和前面一段关于康德的引文,在根据原文翻译时,参考了我的同事余治平研究员提供的德文翻译,在此表示感谢。——译者

② 在最近古耶尔(Guyer)和伍德(Wood)的翻译中是第204—205页。

第 14 章，尽管在其他章节，特别是在第 9 章的“思想流”和第 10 章的“自我意识”中，也有许多相关的讨论和分析。休谟和诸如詹姆斯·密尔(James Mill)之类的追随休谟的英国哲学家的思想的一个核心观点是，通过联想或通过任何其他机制从简单思想来建立复杂思想，詹姆斯对此表示怀疑。他完全拒绝承认，由简单思想合成复杂思想。然而，除此之外，他很赞成联想作为对心灵或(正如他通常提出的那样)大脑的工作机制作出基本的因果性说明的重要性。下面是他在论述联想这一章的最后一页上所说的：

> 上一章，我们已经求助于联想来说明在提高辨别力时的使用效果。下一章，我们会看到，它在其他过程中发挥的绝大部分作用的大量证明，然后，很快会承认，无论如何，任何科学中很少分析原理证明比通常看成是含糊阐述的原理能产生更好的结果。我们自己更明确地阐述这个问题和避免在各种诱因与只知道的关系之间造成混淆的企图，必须使我们重视对未感觉到这种混淆的那些人的大量工作。从实践的观点来看，一个人自我恭维说，当他颠覆了原子思想的理论，或表明，在产生了联想之后，才有思想间的邻近和相似性时，已经沉重打击了联想心理学，这真的是用歪曲对方论点的方法驳倒对方。当你把“思想”一方面转化为“对象”，另一方面转化为“大脑过程”之后，联想心理学的整个体系仍然是站得住脚的；能力分析和操作分析，用这些术语与用在传统意义上使用的那些术语一样是结论性的。
>
> (*Principles*: p. 604)

89

詹姆斯对联想律和相关问题的讨论极其广泛而多变，以致这里无法以一种认真的方式来作出总结。我只提到与我的下列

同构论题相关的重要观点：同构是表征的中心概念的精华。首先，詹姆斯不仅在这一章，而且在许多其他地方，都谈到了意象和意象性。从这些许多不同段落的语境来看，相当显而易见的是，他几乎总是记得，在产生意象的对象或现象与意象本身之间，大约有一种同构关系。同构概念的连续使用，在对视觉意象心理学的任何一种详细的研究中，好像几乎是无法避免的。詹姆斯在《心理学原理》的头几页（绪言的第6页）里所采纳的关于外部世界的科学实在论的态度，也支持了这一点。同时，重要的是强调，关于心理意象的这些文献，不管是非正式的文献，还是系统的文献，几乎从来都没有任何严格形式的或定量的企图，来描述意象和产生意象的那个对象之间的关系，或者正如在休谟的情形中那样，连续知觉或意象之间的关系。正如我们从投影几何和透视理论中已经知道的那样，对在物体的表征方面的某些必要约束（例如，通过双眼视力）的精确描述，是错综复杂的。研究这个主题是一门专门的学科，通常不是对意象或心理表征的一般讨论的部分。第6章的4—7节，即关于视觉空间的那几节，在某些数学细节方面，讨论了相关的但不是准确对应的透视问题的形式发展。

关于詹姆斯对联想的心理学现象的分析与评论，我总结了六个方面。对于每个方面而言，我指出了所引证的标准版本的相关页码。

第一，詹姆斯在几个地方提出的观点是（第554页）：联想不在思想之间，而在所思考的事物之间。尽管他不清楚这一点，但他也许承认，联想关系在这些事物的意象之间。他确实把它看成是他全盘否认存在着简单思想的一部分，这些简单思想被看成是心理学思想的原子思想。

第二，对于詹姆斯来说，有许多种联想，而且不管对一般规

律说什么,重要的是承认这一点。他根据“想到的”联结来阐述这一点。

> 决不可能简单地阐述想到的这种错综复杂的联结。每一种可想象的联结都可能是想到的——共存的、连续的、相似的、相反的、矛盾的、原因与结果、手段与目的,属与种、部分与整体、实体与属性、早与晚、大与小、地主与租户、主人与仆人——天知道有什么,因为这个列表真的是无止境的。
>
> (*Principles*: p. 551)

90

第三,联想的一个重要的基本规律是邻近(特别是在时间中的邻近,要么同时,要么前后相继)的规律(pp. 563 - 566)。此外,第四,针对联想的这种基本规律,即邻近律,他确实根据神经系统中的习惯性的规律给出了一种定性的说明。这里是对冗长讨论的一种总结:

> 通过对象在思想中或经验中以前的邻近性想到的对象的心理学的联想律,在心灵里就是下列物理事实的一种效果:神经流最容易通过那些已经使用最多的传导来传播自身。
>
> (*Principles*: p. 563)

在第2章中,詹姆斯在这段引文及其相伴随的讨论中所说的话带有现代的气质,而且是从他的时代已知的大脑活动的详尽论述中得出的结论。另一方面,这种讨论是如此地定性,以致从实际上提出的对联想的科学说明来看,几乎没有多大新意,当然,休谟至少在他的时代的概念框架内认为,这是不可能的。

第五(这一点对我们这里的讨论是重要的),詹姆斯否认,存在着类似联想的根本规律或基本规律,我把它看作类似于休谟的类似性概念(p. 578ff)。他根据合成的思想或想法表达关于

联想的某些肯定方面。到这一章快要结束时，他有了关于实际上不重要的相似性的很明确的下列陈述。

> 当把我们从一个对象或一种思想带到另一个对象或另一种思想时，在两个对象之间或两种思想之间的相似性（假如思想之间存在着相似性的话），根本没有任何诱因。这只不过是一种结果——当这种情况碰巧以某种特殊的和可归属的方式起作用时，通常的因果动因的效果。但是，普通作者议论说，好像对象的相似性本身就是一种原因，与习惯相并列，并独立于习惯，但像习惯一样能推进对对象的深思。这是相当不可理解的。两种东西的相似性，只有当两者都在那里时，才会存在——不管是在物理学领域，还是在超自然现象的领域，把相似性说成是任何事物产生的原因，是无意义的。
>
> （*Principles*：p. 591）

然而，与这一段相反，他确实对这一章前面的全等概念发表了肯定看法：

> 于是，一种即逝思想的有趣部分为什么应该唤醒这种表征而不是那种表征的所有理由是**习惯、近因、生动逼真和情感一致**。我们可以实话实说，**在大多数情况下，即将出现的表征，要么是习惯的、新近的，要么是逼真的和将是一致的**……然而，尽管事实是，一系列表征因此而是从完全的非决定论中拯救出来的，并限于少数几个类，这些类的特征性质是通过我们过去体验的本性来确定的，但是，我们仍然必须承认，在我们表征的链条中，极其大量的术语还是落在了所有指定规则的范围之外。
>
> （*Principles*：p. 577）

在描述他所意指的情感一致是什么意思时,他根据我们想到的 91
意象和我们所处的情感状态之间的关系,举出了生动逼真的例子。这样,例如,如果我们是快乐的,我们就不会想到郁闷的意象。另一方面,如果我们的心情处于忧郁的或悲伤的状态,那么我们将有一个消极的和悲观类型的意象。

詹姆斯对这些问题的思考是错综复杂的。正如上一段所表明的那样,他没有考虑到,相似性可能是联想的基础。但是,他在这一章的最后一节赞成地引用了亚历山大·贝恩(Alexander Bain)(他漂亮地举出了这个详细的例子)的一长段。正如水磨和后来的蒸汽机例示的那样,新型能源或动力的发现者,一定是根据用途,而不是根据物理特性,进行联想,因为,例如,物理特性与水磨或蒸汽机是完全不同的,而且蒸汽机本身在物理上与水磨也是完全不同的。在引用这类例子时,他也说了一些类似于对情感一致所说的话。就像上一段论述多种联想时一样,詹姆斯依据他的多元论的态度,似乎愿意承认,只要在两种现象或两种东西之间存在着用普通术语来说是相似的一种性质,那么正如我所指出的那样,这种同构形式就能成为一种联想的基础。这种特性不必是所研究的对象或现象的一种明确的物理特性,乃至一种明确的物理关系,而可能是由意向和用途产生的某种东西。当我们这样考虑问题时,我们就能够进一步回顾并承认,邻近原理本身是一种特殊的同构形式,因为共享了空间和时间。但是,正如詹姆斯指出的那样,这并不是说,我们同等地把一切都与任何别的东西联系起来,而是甚至在我们通过邻近进行联想的情况下,我们也有很大的选择性。

詹姆斯对许多事例的零乱而大范围的讨论,使得类似性或相似性的各种概念,在被康德忽视之后又重新发挥了作用。但是,他没有以系统的和形式的方式这样做。毫无疑问,在某种意

义上，类似性、相似性或同构，在从休谟转到詹姆斯时，已经失去了阵地，或多或少独立于康德的情形中所发生的情况。

关于詹姆斯的最后一点。在这一章即将结束时，他在自发的联想(spontaneous association)和他所说的随意的联想(voluntary association)(即我们当前所说的意向性联想)之间作出了一个重要的区分(pp. 583 - 588)。自发的或自由的联想通常是这些讨论的焦点，但正如詹姆斯指出的，像在忘了某人名字的情形中那样，在回忆时，特别是，当如果没有某种意向性的努力，这种回忆就不会轻易出现时，集中的或意向性的联想是非常重要的机制。联想的另一个重要的意向用法是在解决问题时，需要用新的联想，集中精力找到解决方法。詹姆斯在这种语境中正确地提到了联想在数学和科学思想中的重要性，作为一个主题还没有得到系统的研究。我在这里将不展开这个论题。它是由詹姆斯涉及到，但显然，这样的意向性联想利用了连贯的或一致的一系列联想，来把解决一个特定问题所需要的概念放在一起。这里，我们再一次使同构概念远离了对象或现象的视觉特性或物理特性，转向使用它们的方式之间的联想。

92

意象的特殊案例。在哲学和心理学领域内，自詹姆斯的《心理学原理》出版以来的一百多年中，现代文献非常之多且各式各样。我不想对这些文献的大多数问题作出任何概述，这里只集中于心理意象性，特别是致力于比较近些年来哲学家一直在谈论的问题与心理学家在他们的实验和理论分析中所关注的那类问题，它们恰好相反。

首先，也许最令人惊奇的结论是，关于意象性的讨论极其罕见，尽管作为一个论题，在关注心理活动的许多哲学家的著作中，其哲学史很悠久。一个好的例子是由皮科克给出的(Peacocke, 1983)。集中于意义与内容的这本认真的和复杂的

著作，在意象性下面还没有一个索引项。缺乏明显的参考并不意味着皮科克不在讨论通常看成是知觉的问题，或者用更加常识的术语来说，人们正在看到的一个物体的心理意象来谈论。在许多方面，他希望用谈论内容替代关于意象性的评论。正如他喜欢指出的那样，谈论内容就是完成所谓的下列短语这种事情："表面上看，似乎这个主题……"(Peacocke, 1983: p. 8)他不想在任何直接的意义上把内容与意象性联系起来，这一点能够从他意指的必须有两个特征的表征内容的描述中看得出来。

> 第一个特征是，表征内容关注体验者的外部世界，因此作出真假评价。第二个特征是，这种内容是内在于经验本身的某种东西——任何一种没有把世界向主体表现为以这种内容所规定的方式存在着的经验，在现象学的意义上是不同的，即一种不同类型的经验。与这两个特征相当一致的是，过去的经验和学习导致出现了具有给定表征内容的经验。
>
> (Peacocke, 1983: pp. 9 – 10)[①]

皮科克对在感觉、知觉和判断——运用他主要回避的过时术语——之间的区分，有些是传统的，有些是新的，说出许多令人感兴趣的东西。正如他以许多不同方式和在许多不同的地方所指出的那样，他的表征内容的概念最确定的是不同于感觉。尽管事实是，他集中于视觉表征，并举了许多熟悉的事例，但是，他不仅避免使用"意象"这个术语，而更重要的是，回避对表征内容与外部世界中的对象的物理特性之间的任何一种同构作出任何断言。此外，为了对我所断言的皮科克的分析不产生误解，这不

① 原文是 Peacock，这个标注有误，应该为 Peacocke。——译者

是说，他避免谈及物理特性，而是正如从上面的引文中所看出的那样，他主要希望把这些特性简单地称为对象的特性，没有在系统的心灵中的视觉意象与看到的对象之间作出区分。我下面所93 要评论的许多心理学文献的观点恰好是，只集中于心灵中的这些视觉意象的重要特性，以及它们如何通过各种不同类型的同构系统地与对象本身联系起来。

偶尔熟悉休谟和更早的传统文献的一位局外人，当他探讨像皮科克的分析一样的现代的认真研究时，可能会认为，现在用投射式变换及其不变性的精致的形式语言，以把一个有必要深化的、系统的元素带到早期的哲学思想中的方式，能阐明休谟关于类似性的一般概念或更早的亚里士多德的形式概念。尽管皮科克的分析在许多方面是慎重的和严肃的，但是，这种努力是达不到的。从认识的、语言的和语义的区别方面来看，他关心的是其他方面的问题。也许，最令人吃惊的证据是，在他的标题为“空间内容和约束”的第 3 章中，皮科克回避了空间意象性和把这些意象与物体的几何方面联系起来的复杂问题，这一章显著的是只有最微不足道的几种几何区别。他只涉及更古老的、更基本的心理学文献，没有涉及现代文献，这是皮科克的观点的特点。令人感兴趣的是，知道他如何看待罗杰·谢泼德(Roger Shepard, 1960)运用意象性的最迷人的事例之一的分析。在这个实验中，要求一个人说出他家有多少个窗户。通常，一个人不会立刻知道这个数目，而且通常完成这项任务的大多数受试者将报告说，他们所做的是，想象考虑每间屋子里有几个窗户，然后，计算窗户的总数。现在，就房子里的每间屋子来说，为了使这种想象起作用，它的背后一定存在着想象的屋子与真实的屋子本身之间的适当的同构概念。当然，我们中的大多数人能够非常漂亮地回答谢泼德的问题。在这个框架中，很重要的理论

观点是,根据我们对熟悉但复杂的物体的心理意象是如何起作用的,来研究我们确实具有什么样的同构关系和没有什么样的同构关系。

从所做的大量实验和分析来看,在心理学中,关于意象性的争论已经持续了一段时间。最近,许多哲学家也精神焕发地参与到这种争论中来。剖析论证双方的一个好的例子来自罗利斯(Rollins, 1989)。他以一种合理的方式为某种同构,即更多地得到相信意象的那些人支持的他所谓的"普通的相似性"(nomic similarity),作出了论证。但是,他也通过描述主义者总结了反对这种同构的论证,描述主义者坚持认为,认知内容不会在对象的意象中被发现,而是在对对象的描述中被发现,即描述对象的语句是最好的表达。福多和佩利夏恩在他们的著名文章中(Fodor and Pylyshyn, 1988)支持了这种观点。迈克尔·泰在他的书中,也对这种争论的这些哲学方面和其他哲学方面作出了很好的评论(Michael Tye, 1991)。另一方面,其实通常对争论双方都很公道的这个文献,显然没有详细地考虑,真正的同构、相似性或类似性的理论看起来像什么的问题。换言之,重要的细节根本没有从一种几何的或结构的观点展开阐述。

意象性的心理学观点。这样,已经强调了在哲学文献中缺少细节的这点——缺少的细节在某种意义上在休谟或詹姆斯的著作中并不缺少——我最后转向芬克(Finke)对心理学文献的好的评论(1989),芬克概述了关于意象的极其大量的心理学实验,而且也采取了一种公平的态度对待在关于意象的争论中产生的许多错综复杂的论证和反驳。芬克围绕五个原理整理了他对心理意象的心理学文献的分析,他为每个原理都提供了支持的实验和分析,还有相应的批评。 94

第一个是关于运用心理意象的信息检索原理。在这里,关

于记忆的实验占有中心阶段。第二个是关于心理意象的视觉特性。正如将要得到辩护或受到批评的原理那样,他陈述了一个知觉等价性原理:

> 当把对象或事件想象成是,实际上感觉到了同样的对象或事件时,视觉系统中的相似机制得到了激活,在这种程度上,意象在功能上等价于知觉。
>
> (*Principles of Mental Imagery*: p. 41)

非常容易地说出一些哲学家的名字,他们会一直强烈反对这种建议:这个原理确实是人类心理意象的真相。

第三个原理是关于心理意象性的空间特性。在为心理意象贴上了空间等价性原理的标签之后,芬克给出了下列阐述:

> 心理意象的元素的空间安排,对应于在实际的物质表面或实际的物理空间中安排对象或它们的部分的方式。
>
> (*Principles of Mental Imagery*: p. 61)

很容易想到关于这个原理的反对意见,但是,它也有能够得到日常经验和大量不同的心理学实验辩护的一种重要吸引力。或许,最认同这个原理的心理学家是斯蒂芬·克斯林(Stephen Kosslyn),他在许多不同的场合对它作出了强烈的辩护(1973、1975、1976、1978、1980)。

第四个原理是关于心理意象的转化,特别是与罗杰·谢泼德的实验和他关于意象的心理旋转的联想联系在一起,例如,在谢泼德和梅茨勒(Metzler)的文章中(1971)。谢泼德和库珀(Cooper)全面地描述了较早的实验和其他许多细节。芬克对转化等价性原理提出了如下的阐述:

> 意象的转化和物理转化呈现出相应的动力学特征,而

且，受同样的运动定律支配。

(*Principles of Mental Imagery*: p. 93)

尽管谢泼德的实验肯定证明了有点像产生了常识的心理旋转概念，但是，对于实际上并不坚信意象的实验心理学家来说，这种明显的形式和芬克陈述这个原理的强同构的假定，很像是把生肉扔给狮子——他们迫不及待地消化它，并用实验表明，它错在哪里。涉猎意象争论的任何一个人，不管是在哲学意义上，还是在心理学意义上，都会知道反对者在哪里摧毁像芬克所阐述的那些原理。他们将提出解释和语境影响在不同情境中发现的一致性的许多方式。因此，在开始毫不保留地接受芬克所作的阐述时，至少需要陈述一些其他条件不变的强条件。另一方面，为这些原理辩护的那些人更多地站在他们的立场上。在没有这样 95 的近似同构的前提下提供一个合理理论的困难，要比许多批评者认识到的许多困难更严峻。

最后，也许最有争议的是，作为他的五个原理中的最后一个原理，芬克陈述了结构等价性原理。“心理意象的结构对应于实际感知体的结构。”(p. 120)

这个原理的思想是，结构关系，即复杂的几何图形部分间的关系，通常被发现保留在心理意象中。正如芬克所指出的那样，关于心理意象的普遍误解之一是，意象是突然或立即产生的，而不是随时间在某些方面建构起来的。(我必须说，由部分建构起来的想法与记忆如何被建构的思想近似得多。许多研究表明，记忆不只是像一台数字计算机那样存取数据，而是通过联想的过程一步一步建构起来的。)

克斯林、赖泽(Reiser)、法拉赫(Farah)和弗利格尔(Fliegel)通过做实验来揭示结构是如何发挥作用的。向受试者描绘由简单的几何形式画出的图形，例如，正方形、长方形和三角形。向

受试者提供不同的描述，这些描述随着描绘这个图形所用的部分的总数的变化而变化。例如，两个长方形，一个是垂直的，另一个是水平的，围绕它们的中点居中对齐，能够被描述成是五个正方形或两个重叠的长方形。向受试者提供两种描述：一种描述是部分的总数最少，另一种描述是部分的总数最多。然后，命令受试者按照这种描绘形成一个已知图形的一种意象。记录下他们形成意象的时间。在一种显著的程度上，形成意象所需要的心理时间与描绘图形时所用的部分的总数之间，存在着一种很好的线性关系。当然，这类问题在很大程度上使人联想到半个多世纪之前的格式塔心理学实验。我发现，这令人震惊的是，如果根据描述主义者的替代说明，或者，相信对心理活动作出完全类似于命题说明的那些人的替代说明，大胆地怀疑这些实验中所用的意象的所有方面。不过，这不是把这种争论引向深入的地方。

我在本章的最后一节一直试图辩护的是，结构的某种相同性概念或同构，不仅在简单的数学事例中，比如在测量理论中或在几何学中，而且在与心理意象一样复杂和丰富的某些方面，都是重要的(即使是有争议的)。在这个案例中，这种复杂性，在哲学和心理学领域内，围绕心理意象的地位至今仍在继续的争论的复杂性中，深入地反映出来。我已经清楚地表达了我自己的同情心，但我承认，意象的批评者们所做的工作，非常有助于深化实验和对这些实验的思考。①

---

① 更深层次的复杂性有许多。当在数学或科学研究或教学中运用视觉图形或草图时，我这里甚至在最表面的层次上都不想考虑的心理学特征，起到了关键的作用。我这里只论及在解决问题的活动中能够观察到的眼动。它们更多地告诉我们，在扫描图形和在必定有限的工作记忆中储存意象或描述之间，连续的和迅速的相互作用。关于这些问题的详尽讨论，参见 Epelboim and Suppes(2001)。动物和人类解决问题的经常易变的本性，至今还没有在关于心理意象的争论中充分反映出来。

4. 不变性

与表征思想复杂地联系在一起的是表征的不变性思想。这一章,我首先在第一节探索不变性与对称性之间的密切关系,像在视觉设计中反映的那样。即使是最简单的设计,我们也早已能够通过计算它们的基本对称群来增进我们的理解,正如我希望用适当的例子和定义来解释的那样。从这些直观的思想很容易转移到更加形式化的不变性、对称性和含义的思想。 97

第二节考察定性的视知觉的不变性问题,正如日常语言特别是英语中的介词所表达的那样。第一、二节在精神上是相对非形式的和直观的。在第三节,我返回到更加形式的但却是基本的问题:从第 3 章的表征定理的观点分析等区间测量理论的不变性问题。在第四节,考察所缺少的基本物理学方程的不变性,然后,在第五节,即最后一节,略述了熵的极好结果,即熵在许多各态历经过程中是完全不变的。

4.1 不变性、对称性和含义

不变性的日常或普遍含义(meaning)定性地提供了对其技术含义的很好的领悟。如果某种东西是不可变更的、不能改变

的或恒定的，那么它就是不变的。当然，由此带来的第一个问题是："什么是不可变更的或不能改变的呢？"直观的回答是，一个对象或对象的集合的特性、或更一般地说某种现象是恒定的，这种回答一般说来与我后面将要考察的技术性的回答也是相当匹配的。常见的例子是一个杯子的形状，当我在周围移动这个杯子、旋转它、把它倒过来等时，这个杯子的形状是不变的，它的重量也是不变的。一个重要的物理学例子是，与经典物理学中熟悉的速度的简单相加相反，在真空中，光速相对于我们测量它的任何一个惯性参照系都是不变的。

表达恒定性的不变性思想是尤其重要的。因此，我们说，在心理学的意义上，某人的态度是不变的或不能改变的，恰好因为这些态度多年来一直是恒定的。当然，我陈述的所有含义，彼此之间大约是同义词。恒定的某种东西是不可更改的。恒定的某种东西是不能改变的，等等。

98

正如许多关于不变性的评论者已经强调的那样，与不变性密切相联系的是对称性。我们最亲密的例子是，人体的两侧大约是对称的。人们有时好像认为，在设计中和建筑结构中，两边对称的心理要求，是我们自己的两侧对称性的一种结果，尽管我这里并不这么认为。很难确定这种心理要求是否是真的，但结构和设计中的对称要求是十分明显的。我这里恰好呈现两个图形。一个是中国人设计的(Jones, 1867，盘子 XIX)。这个图形的垂直平分线是两边对称的轴线。第二个图形是帕拉迪奥(Palladio)(1570/1965)提供的一个别墅的设计图和正视图。这个设计图和正视图的图样甚至比结构本身都更明显地努力使两边尽可能地对称。

我们能够从两边对称的特殊案例概括出源于费利克斯·克莱因(Felix Klein)的几何观点。这种观点能以这样的方式来陈



图 1　是中国人设计的瓷盘，大概在 19 世纪

述。“为一个空间变换的给定群寻找其不变的特性、量或关系。”从这里阐述的观点来看，因为一个给定的几何意指由基本集合、组成的空间点和在这个集合上的关系与函数组成的一个集合论结构，所以，把一种几何结构转换为在相同的空间与它同构的另一种几何结构的一种变换，是这个空间的一种**自同构**(automorphism)。克莱因的观点是，一种几何的自同构群提供了定义这种几何的另一种方式。或者，以更纯粹的克莱因的方式来说，一个空间变换的任何一个自同构群都定义了一种几何，而且，存在着在自同构群的条件下不变的这种几何的关系与函数。

99

下面是来自克莱因的一段著名引文，取自他 1872 年在爱尔兰根的演讲(1893)。我在英语翻译时偶尔做了很小的改动。

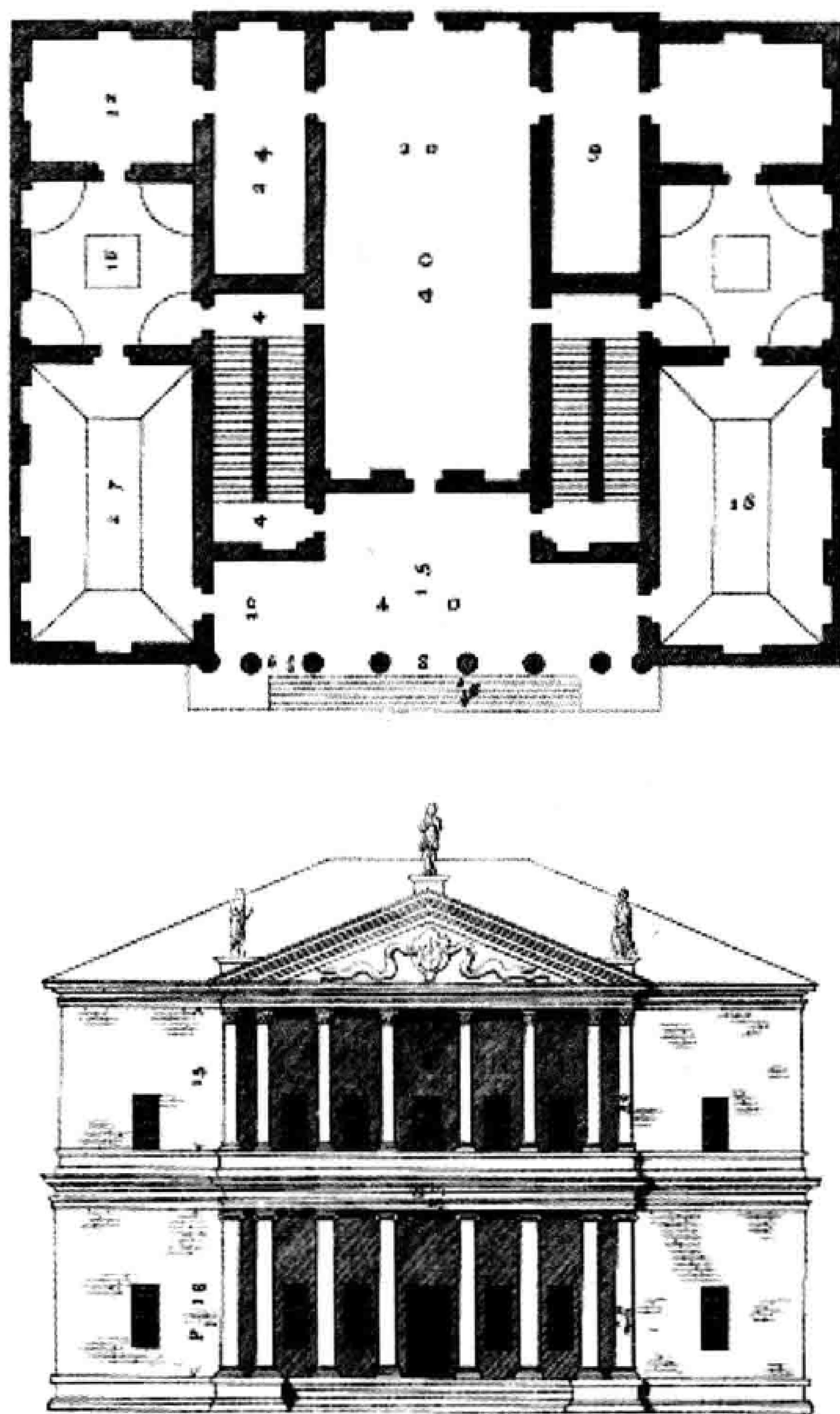


图 2　帕拉迪奥别墅的设计图和正视图

因为几何特性从其思想本身来看独立于所考虑的位型在空间占有的位置，独立于它的绝对量，最后，在某种意义上，也独立于排列它的部分。因此，一个位型的特性，通过

空间的任何运动、通过变换到相似的位型、通过变换到关于一个平面对称的位型(偏转),还有通过这些变换的任何一种结合,仍然是不变的。我们把所有这些变换的总数命名为空间变换的**主群:主群的变换不改变几何特性**。而且相反,**在主群变换下,几何特性通过它们保持的不变性来描述**。因为,如果我们把目前的空间看成是不动的,等等,看成是一种严格的流形,那么每个图形都有一个个体特征;在它作为个体拥有的所有特性中,只有纯几何的特性在主群变换下保持不变。 100

(Klein: p.218)

在这个框架内,我们说,一个几何图形在给定群的条件下是**对称的**,如果这个群的每个自同构都把这个图形映射到它自身的话。例如,在平面图形的情况下,沿着我们所谓的一条常见的对称线,通过一种镜像反射,这样就映射出一个两边对称的图形。在各种不同旋转条件下对称的图形,也是对称性的常见例子,首先圆是如此,但当然,正方形在围绕其中心任意旋转90度,也是对称的。注意这种区别。正方形有一个旋转的很小的有穷群,在这个有穷群下,它是不变的,而圆在一个旋转的无穷群下是不变的。

同样,像球和立方体之类的三维图形,在移动或变换的各种群下也是不变的。在这里,理想化是重要的。在旋转或偏转下,大多数真正的物体,在某一个可观察的方面,即使是瞬间的,看起来也是不同的,或者就是不同的。

几何不变性的熟悉例子是欧几里得几何的全等关系和介中性(betweenness)关系,只根据这两个概念能够把欧几里得几何公理化。

为什么对称性和群会很自然地联系到一起,是值得评论的。

我们可以这样来考虑。如果 φ 把 F 映射到自身，那么在一个空间 A 中，一个几何图形 F 在把 A 映射到 A 的单一变换 φ 的条件下是**对称的**。现在，如果这一点成立的话，逆映射 φ^{-1} 应该同样如此，因为 $\varphi^{-1}\circ\varphi$ 恰好是 A 的恒等映射。此外，自然会料到，这个对称特性是闭包的。如果 φ_1 和 φ_2 两者都把 F 映射到自身，那么它们的合成 $\varphi_1\circ\varphi_2$ 也应该如此。变换集合 G 的这两种特性：

(i) 如果 $\varphi\in G$，那么 $\varphi^{-1}\in G$，

(ii) 如果 φ_1，$\varphi_2\in G$，那么 $\varphi_1\circ\varphi_2\in G$，

足以保证 G 是给定空间 A 的一个变换群或自同构，因为变换的合成必定是结合的，而且 $\varphi\circ\varphi^{-1}$ 是恒等式。这个相同的论证也表明，为什么不变性和群是如此自然地不可分开。①

当艺术家或建筑师绘制一个对称图形或进行一项设计时，无疑，通常的情形是，他们明显地不会根据潜在地内嵌了图形或设计的空间的自同构来考虑问题。即使是古希腊时期老练的数学家也没有一个明显的群概念，但一旦揭示出群概念，它似乎就是直观的对称性思想的一种自然的形式化。

几千年来，人们一直认可设计中的各种类型的对称性。图3所示的有两条对角线的正方形看起来是相同的，也就是说，它在平面上反转90°、180°、270°，当然还有360°的条件下，外表是不变的。（当然，以哪种方式旋转是一个约定问题，我多少遵守了标准。）

101

---

① 2.3节已经给出了群的一般公理，不只是变换群，但正如在3.3节一开始所提到的那样，每个群都同构于一个变换群，这是把变换群看成是群的最直观模型的一个理由。关于群的抽象概念的详细发展史，可参见武辛的著作（Wussing，1984）。

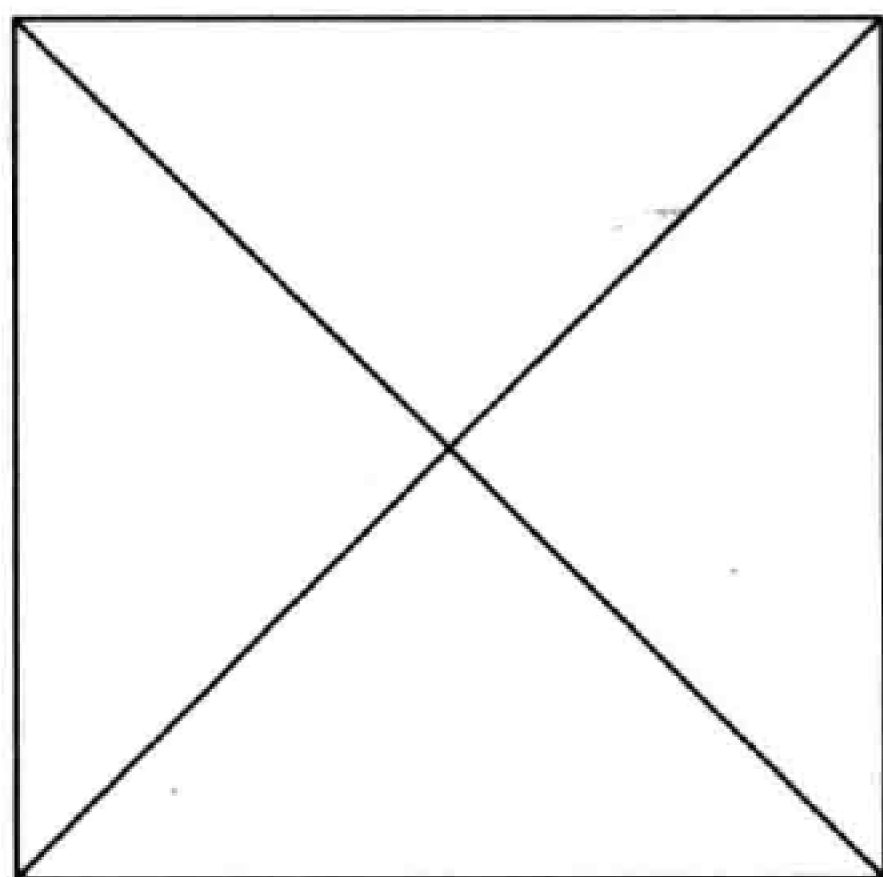


图 3　有两条对角线的正方形

另一方面，图 4 所示的有一条对角线的正方形，只有在围绕其中心反转 180°和 360°的条件下，外表才是不变的。

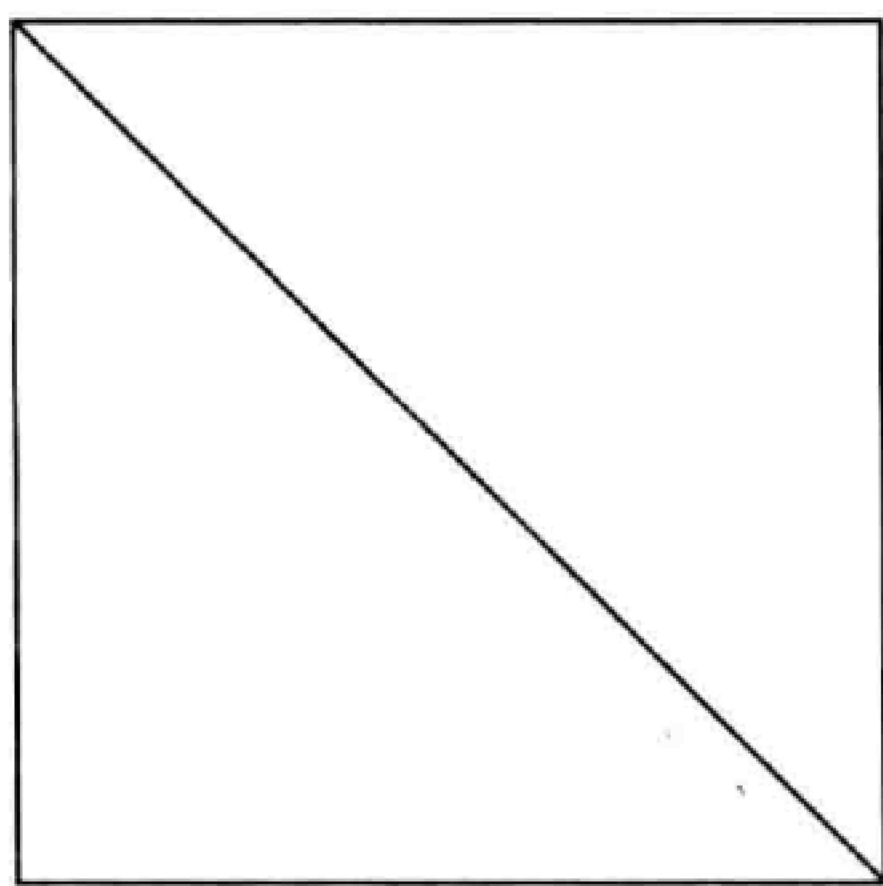


图 4　有一条对角线的正方形

我们只从形式上进行适度的增减，就能够定义这两个图形在反向旋转条件下呈现出的对称性的两个群。在这里，**群**是在上面和第 2 章定义的技术性的数学意义上进行解释。因为让我们首先考虑图 4 是更简单的。就符号而言，设 $r(n)$是一个 n 度的反转。首先注意，$r(0) = r(360)$，而且，$r(0)$是这个群的单位　102

元素。全集显然恰好是{$r(0)$, $r(180)$},而且,这个群的二元运算∘是接下来的 $r(m)$随另一个 $r(n)$的一种反转的运算,我们写为: $r(n)\circ r(m)$。一旦注意到 $r(180)$是自身的反转,当然,$r(0)$也是如此,读者就能轻易地证明,满足了 2.3 节的群公理。

在图 3 的情形中,反转对称群较大,是集合{$r(0)$, $r(90)$, $r(180)$, $r(270)$},而且,在这里,除了 $r(0)$之外,每个"和"的反转都是 360。例如,$r(270)$的反转是 $r(90)$,因为

$$r(90)\circ r(270)=r(0)$$

在规则图形特别是二维图形中的对称性的全面发展是由托特(Toth)提供的(1964)。这是含有丰富的数学史和艺术史的一个主题,但我把它搁在一边,考虑与描述科学理论的结构更直接相关的概念与结果。

关于不变性的一般逻辑结果。一个经典的结果很好地表达了下列重要思想:相对于世界的任何一个一一对应的映射,都是不变的,或者,对所考虑的任何一个模型都是普遍的,这恰好是逻辑关系。这是对几何的克莱因变换进路和群论进路的最有可能的概括。让我来给出两种数学表述,一种表述是回到廷登鲍姆(Tindenbaum)和塔斯基的早期著作(1934 - 1935/1983, p. 385):

> 大致说来,定理 1 陈述,能够通过纯逻辑方法表达的对象(个体、类和关系,等等)之间的每种关系,相对于"世界"(即所有个体的类)的一一对应地映射到自身,都是不变的,而且,这种不变性在逻辑上是可证明的。这个定理在特定的直观考虑中,当然是合理的,也已经作为一个前提来使用。然而,以前它从来没有得到精确的阐述和准确的证明。

一个密切相关但不同的数学表述,在很久之后,才由塔斯基

和吉万特(Givant)给出(1987, p. 57):

> (i) 给定一个基本的全域U,任何一个派生的全域$U\approx$[例如,笛卡儿积$U\times U$]的一个项M,如果在U的每一种排列P的条件下,都是不变的,那么就被说成是逻辑的或一个逻辑对象。基于(i),人们能表明,例如,对于每个(非空的)U来说,在U的元素之间只有四种逻辑二元关系:全域关系$U\times U$、空关系$\varnothing$、恒等关系……,以及相异关系……

后面的这个表述所揭示的是,只有四种逻辑的二元关系。在这个全域的任意排列的条件下,没有一个其他的二元关系是不变的,即是恒定的。

含义。很容易从对称性过渡到含义。这里是一些熟悉的例子,像亚里士多德的空间的上下这样的方向在欧几里得几何中 103 是没有意义的。① 为什么呢? 因为在欧几里得的运动群的条件下,即旋转、平移和偏转的条件下,它是可变的。

在欧几里得几何中,有意义的是,比较两条线段的长度,不管它们是否平行。然而,在仿射几何中,为了使这样一种比较有意义,两条线段必须是平行的,包括它们位于一条相同的直线上。事实上,仿射几何最显著的特征是,它不是一个度量空间。当我们从一条线到另一条线改变方向时,沿着这两条线,根本不存在共同的测量。只有对于平行的线段来说,才存在着一个有意义的全等概念,这是当我们谈到仿射全等时,我们的意思所在。

---

① 下面是引自亚里士多德《论天》(*On the Heavens*)中很有名的一段。

> 存在着某些东西,它们的本性总是向着远离中心的方向运动,而其他总是向着中心运动。我把第一个说成是向上运动,第二个说成是向下运动。有些人否认,在世界上存在着上或下,但这是不合理的。
>
> (Aristotle, 308a14 - 17, Guthrie 翻译,1960)

在投影几何中，在从给定的透视点把一条线投影到另一条线的条件下，甚至介中性的仿射概念都得不到保持。在这种情况下，我们必须转到四个位置之间的关系，即分离关系：$ab\ S\ cd$，当且仅当，a 和 b 与 c 和 d 是分离的，反之亦然。如图 5 所示，如何在投影条件下保持分离，但不保持介中性。例如，我们能很容易地看出，在这个图形中，b 位于 a 和 d 之间，但投影 b' 并不位于 a' 和 d' 之间。另一方面，对 ab 与对 cd 是分离的，反之亦然；此外，在投影条件下，对 $a'b'$ 与对 $c'd'$ 是分离的，反之亦然。

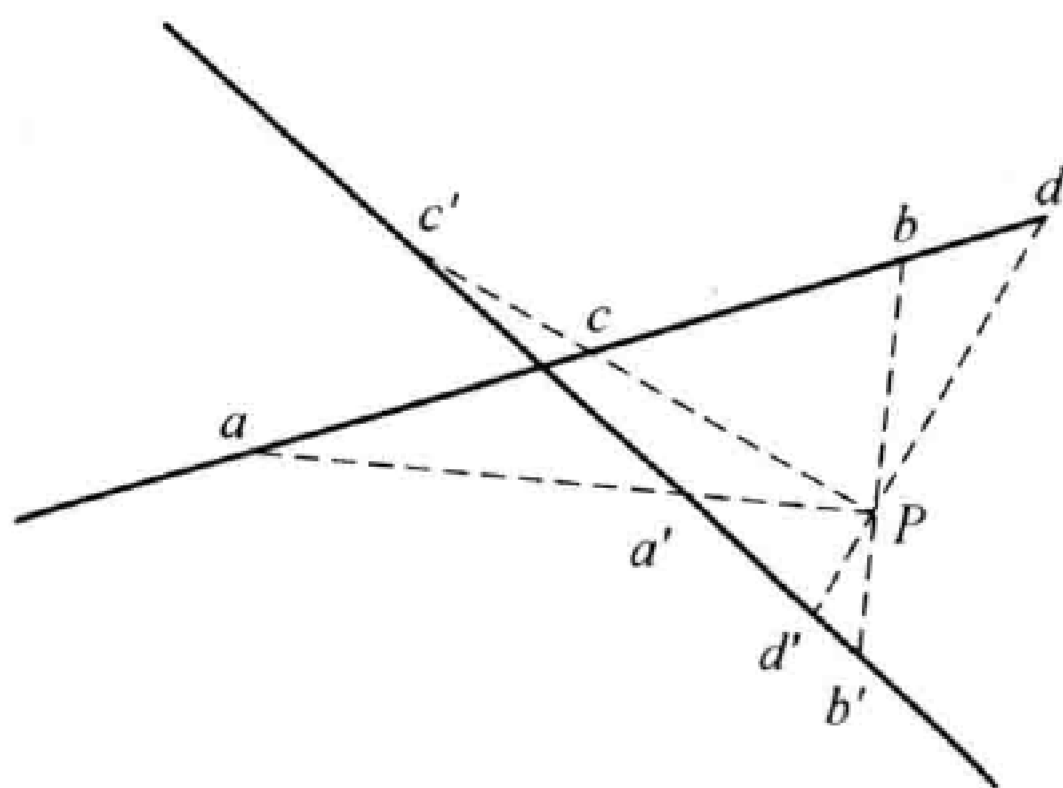


图 5　这个图表明，在从给定的透视点 *P* 把一条线投影到另一条线的条件下，介中性的可变性和分离的不变性。

物理学中的客观含义。在物理学中，存在着的一种倾向是，超越简单地谈论意义性（meaningfulness），而谈论客观含义（objective meaning）。对于彼此以不变的速度作相对运动的观察者来说，重要的概念是不变关系的概念，或更一般地说，是不变性概念。这样，例如，在经典物理学中，彼此作相对运动的两位惯性的观察者应该确证，所进行的相同的测量真的是不变的。

104　这并不意味着，他们在自己的参照系中，用他们各自的测量程序

作不出正确的测量。[1] 但正如我们所知道的那样，比如说，方向在一般情况下对这两位观察者来说会是不同的，其他特性也是如此。另一方面，我们期望坚持，对于所有的观察者来说，用他们的相同刻度的测量仪器，同时测量空间两点之间的距离，他们测量出的数值是相同的。时间也一样。对于所有惯性的经典观察者来说，用具有相同刻度的时钟测量时间间隔，他们测量出的时值是相同的。正如在经典物理学的基本讨论中所强调的那样，空间和时间间隔与观察者无关，但是，宇宙中心的概念、时间的开端，或者，处于绝对静止的状态，是可变的。它们没有任何物理意义，尽管较早的观点，比如，亚里士多德物理学中的那些观点，与此相反。当然，亚里士多德使用的这些概念决不是荒谬的。现象学的经验当然为我们提供了一个绝对静止的自然概念，即相对于地球是静止的自然概念，对于宇宙中心来说，也是一样的。另一方面，更远离熟悉经验的一个问题是，亚里士多德非常令人满意地认为，世界是永恒的，即时间没有开端，为后来的基督教神学家，比如，16世纪的约翰·菲洛泼纳斯(John Philoponus)，提出了问题。

这里是来自狭义相对论的第二个例子。一个运动质点的"固有"时间，对于两位惯性的观察者来说，是不变的，这种不变性足以得出，他们相对于自己参照系所进行的观察，通过洛伦兹变换联系起来。在一个惯性质点轨道上的两个点的固有时间间隔 τ_{12} 满足下列方程：

---

① 在不平行的轨道上彼此相对运动的观察者，测量同一个被观察物体的速度时，所得到的速度值是完全不同的，但这并不意味着他们的观察在通常的认识论意义上是主观的。许多彼此静止的观察者，在主体间的意义上，他们的测量是完全一致的。这里提出的客观含义概念，在物理学中，而不是在含义的一般哲学讨论中，是共同的。在科学哲学的范围内，这是相对于一个理论定义有意义的一种好的方式，即相对于这个理论，客观的含义是不变的。

$$\tau_{12}^2 = (t_1 - t_2)^2 - \frac{1}{c^2}[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2]$$

这里的重要概念是,固有时间具有独立于观察者的特殊参照系的客观含义。此外,固有时间具有完全不变的重要特征,因为它的不变性决定了洛伦兹变换群(Suppes, 1959)。第 6 章提供有关这两个事例的更多细节。

第三个例子是纳脱(Noether)的经典定理。大致地说,这个定理指出,力学中的每个守恒定律都对应于一个不变量,它是运动方程的初积分。

事例 A。一个粒子系统的动量守恒意味着,这个系统的质心以恒定的,即不变的速度运动。

事例 B。一个系统绕一条已知的线旋转。那么,关于作为轴的这条线的角动量是不随时间改变的,即是恒定的。

这些各种不同的几何事例和物理学事例清楚地表明,不变性概念不仅与某些变换群相关,而且在更基础和基本的意义上,与某个理论相关。一个理论模型的不变特性是研究不变性的焦点。这样,例如,亚里士多德的天体理论很自然地导致这些天体是永恒的,因为它们是不变的,因而时间没有开端。

105

在物理宇宙的托勒密学说、哥白尼学说、伽利略学说、牛顿理论和爱因斯坦的狭义相对论中,空间的均匀性潜在或显在地假设了跨越许多理论具有的一种不变性。阐述这个问题的一种技术性的几何方式是这种基本假设:正如第 6 章所说明的那样,空间与时间是一个有许多不变结果的四维仿射空间。

4.2 定性的视知觉的不变性

在这一节中,普通空间的不同表达方式是根据它们的普通含

义引入的，所要问的问题是：就每一种表达方式都是不变的而言，什么是自然的几何理论呢？在第6章中，针对视觉空间的各种不同的实验结果，也提出了相同的问题，这对视觉空间是欧几里得几何的论点提出了挑战。正如已经暗示的那样，我将根据几何和物理学中的实践，把在一个已知变换群下的不变性与对称性原理联系起来。不变性意味着对称性，而且，对称性意味着不变性。

归入本节标题下面的概念、结果和问题，在过去一直没有得到很认真的研究，但在近二十年左右，受到了更多的关注。我在这种关系中特别提到了鲍尔曼(Bowerman, 1989)、克兰格(Crangle)和苏佩斯(1989)以及勒韦(Levelt, 1982, 1984)的工作。这些出版物中的参考文献为追溯更早的文献提供了很好的向导。我这里想到的一个典型问题是勒韦已经完全研究过的那类问题，例如，为了给出空间方向，运用语言的方式，以及我们提供这个方向或准确地描述视觉场景的能力的局限性。在我自己与克兰格的早期工作的情形中，我们尤其关注分析英语中常见的不同介词的用法所隐含的几何。

我们很容易能把分析结果总结为表1，表1在转引时，对克兰格和苏佩斯(1989)过去的列表作了某些修改。所涉及的几何类型是标准的，也许，定向物理空间的几何例外。例如，在句子"**铅笔在盒子里**"(这里，假定盒子是盖着的)的情形中，显然，只需要纯拓扑的不变性概念。另一方面，在句子"**玛丽正坐在约瑟和玛丽亚之间**"的情形中，很容易看到，几何要求是仿射的。然而，像在"**铅笔在盒子旁边**"这个句子中那样，一旦引入度量的思想，我们就必须从仿射几何转向某个潜在的全等概念，像在欧几里得几何中反映的那样。尽管在这种分析中，我可以更精炼，但注意到，在这个句子中，相当令人满意的是，使用了绝对几何，即它是欧几里得几何减去标准的欧几里得公理。这个公理断言，

给定一个点 a 和不经过点 a 的一条线 L，那么，在由 aL 形成的平面中，通过 a 点最多有一条线与线 L 不相交。我们把这个定理的否定式增加到绝对几何，得到的是双曲几何，而不是欧几里得几何。似乎对我来说，绝对几何的全等关系足以很好地满足通常谈到的接近度概念。事实上，要求相对技术性的几何结果，使我们从绝对几何转向欧几里得几何，例如，在建筑和作图中所要求的那类技术细节。绝对几何适当地表达了欧几里得几何的许多基本结果，注意到这种适当性的另一种方式是，欧几里得《几何原本》第一卷的头 26 个命题，在绝对几何中是可证明的。另一方面，在介词 on 的情况下，显然，要求垂直方向的概念，这是在欧几里得几何中完全缺少的一个概念，事实上，这个概念在欧几里得几何中是不可定义的。在具有一个自然的内在方向的对象的情况下，要求一种不同类型的定向。比如，考虑表 1 中给出的句子：**狗在房子的前面**。最后，在许多过程中，谈论静止的空间几何，是不充分的，但是，为了充分讨论，人们需要空间-时间的假设。表 1 的最后举了一个例子。

106

表 1　几何类型和介词用法的例子

拓扑学	铅笔在盒子里。(盒子是盖着的) 一条绳子与另一条绳子绕在一起。
仿射几何	铅笔在盒子里。(盒子是打开的) 玛丽正坐在约瑟和玛丽亚之间。
定向物理空间的几何	本子在桌子上面。 调节桌子上面的台灯。
投影几何	邮局在山坡上。 杯子在盘子的左边。
包括有定向轴的图形与形状的几何	狗在房子的前面。 铅笔在椅子的后面。
经典时-空的几何	她在厨房里削苹果。

我现在想看的是，研究几何的情况所需要的不是标准的公理类型。只重复拓扑几何、投影几何、仿射几何、绝对几何和欧几里得几何的标准公理，是完全不适当的。对这些几何的最新近的描述，会在苏佩斯等人的著作(1989，第13章)中找到。我所要做的是，参考成为这些不同公理基础的原始概念。

定向物理空间。毋庸置疑，最明显地不满足通常谈论的空间关系的绝对几何或欧几里得几何的问题是，根本没有垂直方向的概念。此外，基于林顿巴姆(Lindenbaum)和塔斯基(1934－1935/1983)关于在欧几里得几何中不可能定义重要的二元关系的事实的著名结果，这个概念在欧几里得几何中甚至是不可定义的。就确定性而言，我想把在一条线上两点之间的介中性的仿射三元关系，以及线段的全等概念作为欧几里得几何的原始概念。然而，在我应该提到的许多方面，用平行和垂直概念更方便，而且，不管怎样，我都将假定，后两个概念是定义明确的。当然，作为欧几里得几何的一种延伸，使垂直概念公理化，有许多不同的方式。一条简单的进路是，在三维的欧几里得几何中，增加垂线的原始集合 ν，然后再增加下列类型的公理。*给定任意一点 a，都存在着通过 a 的一条垂线。如果 K 是一条垂线，而且 L 平行于 K，那么 L 是一条垂线。*

107

然而，这条进路有两个困难。第一个也是概念上最基本的困难是，当我们在日常环境里四处走动时，我们的物理空间的自然概念是，它的方向，在垂直和平行的意义上，是被惟一地固定了的。我们没有任意的垂直的方向变换，也没有平行的方向变换。我们自然有相应程度的东、南、西、北的概念。第二个困难也与这个困难密切相关，人们在讨论物理空间的本性时很早就承认，我们很难把地球表面看成是含有水平线和垂直于水平线的垂线的一个平面。从日常经验来看，定向物理空间的自然几

何——马上要阐述的一种观点——是地球的中心成为一个重要概念的球面几何，这与古代的亚里士多德和托勒密的观点是一样的。

亚里士多德直接把知觉证据用作他关于地球是球形的结论的部分论据（《论天》，第 2 卷，第 14 章，297b）："如果地球不是球形的，月蚀就不可能呈现出它们所是的形状，而且，当我们到达北极或南极时，观察行星不会显示出它们具有的变差。"

托勒密在《天文学大成》中的论据更充分。因为这些论据完全是知觉的，许多读者对此不太熟悉，所以，我全文引用。这段引文引自《天文学大成》第一卷的第四部分，这本书是在亚里士多德的著作出版了四百多年后写成的。

> 同样被看成是一个整体的地球，显然是球形的，从下列考虑中能最好地抓住这一点。我们再一次能看到，对于世上的每个人来说，太阳、月亮等星体不会同时升落，但确实是，越靠近东的那些人，看到太阳升起越早，越靠近西的那些人，看到太阳升起越晚。因为我们发现，[对所有的观察者来说]同时发生的蚀现象，特别是月蚀现象，仍然无法被所有的观察者都记录为是在同一时间（即与月亮等距）发生的。更确切地说，更东方的观察者记录下的时间总是比更西方的观察者记录下的时间晚一些。我们发现，时差正比于[观察的]位置之间的距离。因此，人们能够合理地断定，地球表面是球形的，因为它的均匀弯曲的表面（因此而把它考虑为一个整体时）以一种有规律的方式依次为观察者的每个集合切开[这个天体]。
>
> 如果地球的形状是任何别的样子，正如人们能够从下列论证中看到的那样，这一切将不会发生。如果它是凹面，

108

那么更靠近西方的那些人会首先看到这些星体升起;如果它是一个平面,那么,对于世上的每一个人来说,它们是同时升落的;如果它是三角形,或正方形,或其他多边形,那么,根据相同的论证,对于所有生活在相同平面表面上的那些人来说,它们是同时升落的。然而,显然没有这种事情发生。它也不可能是圆柱形的,因为在东西方向上是曲面,以及接近宇宙的两极较平坦,有人会假定这一点似乎更可信。从下面来看,这是显然的:对于生活在曲面上的那些人来说,没有一个星体会是常能看见的,然而,对于所有的观察者来说,或者所有的星体都会升起和落下,或者同样的星体,因为离每一极的[天体]距离都相等,所以总是看不见。事实上,我们越向北极行进,南极的星体消失越多,而北极的星体出现越多。因此,显然,在这里地球的过分弯曲以一种有规律的方式从南北方向切开了[天体],从而证明在所有方向上[地球]都是球形的。

进一步的考虑是,如果我们在海上开船向着任何一个方向的大山或高处驶去,那么会看到它们在逐渐地变大,好像正在从过去淹没了它们的海平面上浮起:这是由于水平面的弯曲造成的。

(*Almagest*: pp. 40 - 41)

亚里士多德的自然运动概念意味着,重的天体落向地球的中心,托勒密在第一卷的第七部分再一次对此进行了很好的论证。

……所有重物运动(我意指固有的[自然]运动)的方向和轨道总是处处与作用点相切画出的严格的平面成直角。这个事实表明,如果[这些落体]没有被地球表面挡阻的话,

那么它们肯定会到达地球本身的中心，因为通向中心的直线也总是在[半径的]交点和切点上与球面相切的平面成直角。

（*Almagest*：pp. 43 - 44）

关于这个和亚里士多德的相似但不太明确的论证，所重要的是注意到了，垂直或向上的概念是沿着半径延伸至超出地球的表面，而不是与一个给定的水平面垂直的线。因此，感觉到的垂直概念就是穿过地球中心的一条线段。平行概念是与地球表面的垂线相正交的一个平面。重体作自然下落运动，这种知觉证据是对这种观点的最有力的论据。

托勒密也用观察星体和行星的运动来固定东西方向和南北极。我们在日常经验中用像太阳的升落之类的论据来确定东西方向。

关于物理空间的这些知觉论证——当然包括关于引力的知觉论证——把欧几里得的运动群简化为恒等的平凡自同构。这意味着，从普遍的观点来看，不变性概念，在考虑定向物理空间的知觉问题时，没有任何实际意义。另一方面，必须强调，这里是在很强的意义上使用普遍概念的。从通常情况下对**普遍**这个术语的使用方式的观点来看，还有许多不能一概而论的对称性，109 它们是重要的，而且，在对空间中的实际对象进行的知觉分析或物理分析当中，不断地被使用。确实，从日常知觉的观点来看，这是一种很直观的结果。把我们的空间视觉与我们来自引力效应的空间感觉结合起来，在直接经验或知觉的意义上，思考像欧几里得的运动群是物理空间的基本群之类的问题，是完全异常的。事实上，对希腊数学的一种引人注目的抽象，不会使定向成为欧几里得的原始公理的一个组成部分。

另一方面，在考虑由视觉、听觉或触觉数据产生的感知图形

或感知现象的对称性时,知觉中的不变性概念以一种不太普遍的方式频繁出现。

从作为一个整体的空间的观点来看,对于表 1 中用到的介词来说,可以把几何的分类看成是一种**外在**的观点。然而,在知觉意义上,这并不是我们研究问题的方式。在日常经验中,我们首先从上面描述的确定的定向物理空间的框架出发。在这个框架内,我们能够阐述空间关系的一种**内在**的观点,像用日常语言所表达的那样。

作为一个例子,我考察介词 in 的不变性。在这种分析和接下来的分析中,把用 in 所表达的关系限于一对刚体。即使这对于描述其他情境的不变性(例如,水中的糖)是极其重要的,但是,只考虑所熟悉的刚体的情境,也能够表达对所拥护的不变性的基本进路的很好的意义。

设 a 和 b 是刚体,且 a 在 b 中。像在表 1 中那样,刚体 b 有两种情形需要考虑。第一种情形是,b 是一个空心的封闭刚体,a 在这个刚体的里面。为了简化符号和分析,我引入了一个限制性假定: a 是一个球,这样,就能忽略 a 的定向。定义 $I(a, b)$ = 在 b 的空心内 a 的中心点的集合。设 $\Phi(a, b)$是 $I(a, b)$ 到自身的所有变换的集合,这个集合代表 a 相对于 b 的位置的可能变化。例如,如果 a 的中心位置在中空的 b 的一边,那么它能够改变到,即运动到内部的另一边,而不影响断言 a 在 b 中的实际情况。

集合 $\Phi(a, b)$是 a 在 b 中的位置的对称群(在函数合成的运算下),但因为中空的 b 在形状上可能是相当不规则的,所以 $\Phi(a, b)$可能没有欧几里得的等距的标准特性。它确实有预期的不变性特性,即如果 $\varphi \in \Phi(a, b)$,那么有中心 c 的 a,在 b 中的 p 点,当且仅当,有中心 c 的 a,在 b 中的 $\varphi(p)$点。

物体 b 本身应该是在固定的物理空间中的刚体运动。如果 b 是封闭的，那么允许围绕一个水平轴旋转，但是，如果 b 是开着的，那么 b 的可能的刚体运动一定只局限在那些保持垂直的方向上，当然，当 b 应该是刚体运动 ψ 时，为了保持内部关系的不变性，物体 a 必须应该是相同的变换 ψ。显然，a 也可能通过 $\varphi \in \Phi(a, b)$ 进行变换，这样，它的全部变换都可能是保留了 in 的不变性的合成 $\varphi \circ \psi$。

能对表 1 中列出的其他介词的内在不变特性给出相似的分

110 析。它们全部都有不同的对称群，但是，每个群的确切本性是由相关刚体的特殊形状和大小(它们从一个语境到另一个语境可以是变化很大的)决定的。在可计算的意义上，不变性的简单理论没有一个适用于作为用日常语言表达的普通的空间关系，这是不奇怪的，因为这种关系的严格的应用范围太复杂，以至于没有一种简单的统一的几何。

进一步的应用。除了提供空间术语的一种分析之外，在不同语言的比较中也会用到上面概述的这条进路。一个很自然的问题是，存在着一种空间知觉的通用语义，还是不同的语言在本质上有不同的语义。当然，在介词的范围内，有微妙的差异。例如，鲍尔曼(1989)指出，我们在英语中既能说 The cup is on the table，也能说，The fly is on the window，而在德语中，我们用介词 auf 代表在桌子上面，而用 an 代表在垂直的窗户上面。

这样的差异是在预料之中的，尽管研究不同语言的介词在它们的语义范围内如何重叠，是一个相当有趣的问题。学习介词的正确用法是在印欧语系中学习第二种语言最困难的问题之一，这是人所共知的事实。

人们的希望可能是，存在着一种通用的空间知觉，这样，我

们就可以寻找跨语言的几何不变性。但如果我们使几何与个别语言密切相符，那么，原始语言，而不是公理，在拥有一种不同几何含义的意义上，可能是不同的。要问的一系列相关问题可能是，小孩子提供的空间介词的用法有什么顺序。这似乎为仔细地考察在不同语言中有什么变化提供了一个特别好的机会，因为在字面意义上使用每一种语言时，支配空间介词的用法规则是相对具体的和明确的。鲍尔曼(1989)对许多细节作出了出色的讨论，但并没有聚焦于系统的几何问题。

4.3 测量理论中的不变性

可能要问的与一个对象或对象集的任何一个被测量的特性相关的问题是，如何惟一地给测量特性赋值。例如，一个小石子的质量可以用克或磅来测量。一旦选定一个单位，给被测量的质量赋值，就是惟一的。

在℃和℉的情况下，温度的测量有不同的特征。在这里，原点和单位是任意选择的。下列两种测量举例说明了其他形式上不同的测量类型：(1) 概率的测度[①]，这是绝对惟一的，以及(2) 像矿石的硬度之类的物理特性或像智力和种族偏见之类的心理学特性的定序测量。

运用这些不同种类的变换，对本章的主要思想来说是基本的。一个经验假设，或者，事实上是使用数量的任何一种陈述，只有当它的真值，在所包含的数量的适当变换下是不变的时，才具有经验意义。举一个例子来说，假定一位心理学家进行 I. Q.

① measurement 这个词在日常生活和物理学中通常译为“测量”，在概率论中通常译为“测度”，在心理学的语境中，有时也译为“衡量”。——译者

的一种定序测量，并且，他认为，在人的智能的等级排列中，某种新测试 T 的分数 $S(a)$ 具有定序的意义。进一步假定，他能获得的他的受试者的年龄是 $A(a)$。那么，问题是，他应该把下列假设看成是在经验上有意义的吗？

111

假设 1. **对于任何一位受试者 a 和 b，如果 $S(a)/A(a)<S(b)/A(b)$，那么 I. Q. $(a)<$ I. Q. (b)。**

从经验含义的不变性刻画的观点来看，答案是否定的。为了明白这一点，设 I. Q. $(a)\geqslant$ I. Q. (b)，再设 $A(a)=7$、$A(b)=12$、$S(a)=3$ 和 $S(b)=7$。对 I. Q. 的数据不作任何变换，对年龄数据也不作任何变换。但设 ϕ 是从 3 到 6 和从 7 到自身的一种增加变换。于是，我们有

$$\frac{3}{7}<\frac{7}{12}$$

但是

$$\frac{6}{7}\geqslant\frac{7}{12}$$

而且，假设 1 的真值在 ϕ 的条件下不是不变的。

关于一个量的变换特征在经验上有意义的东西是，它以精确的形式体现了，在获得一种给定测量所用的经验运算和相应的算术运算或关系之间，结构的同构如何是惟一的。例如，如果经验运算只是按照某种特征对一个对象集作出排序，那么相应的算术关系就是小于(或大于)关系，而且以一种保持经验顺序的方式把对象映射到数的任何两个函数都是适当的。

这就很容易表明，如果 f_1 和 f_2 在这种意义上是适当的，那么它们是通过一种单调增加的变换联系起来的。在这种情境中，只有在单调增加的变换下是不变的那些算术运算和关系，才

是有经验意义的。

当我们从考察数量转向考察更复杂理论的模型时,我们得到了具有类似特征的结果。例如,在对经典力学的考察中,当确定了测量单位时,我们得到了惟一地符合伽利略变换(即变换到某个其他惯性系的一种变换)的一种表征。在质点力学的相对论结构的情况下,这种惟一性是符合洛伦兹变换的。第 6 章分析这两种变换。

为了举出一个基本的例子,我们能陈述与有穷弱排序的表征定理(定理 3.3.1)相对应的这种惟一性结果。**设 $\mathfrak{A}=(A, \preceq)$ 是一个有穷弱排序。于是,任何两个同态的数值弱排序通过一个严格的递增数值函数联系起来。**换一种语言来说,有穷弱排序的数值表征惟一地符合定序变换。符合定序变换的不变性,不是测量的很有说服力的一种特性,而且,正是由于这种理由,假设 1 原来证明是没有意义的。

返回到问题的一般方案,一个表征定理通常应该伴随着一个相匹配的不变性定理,这个定理主张,一个结构的表征是惟一的。在数学上简单的和直接的情形中,很容易把这个群鉴别为某个众所周知的变换群。对于更复杂的结构来说,例如,满足一个科学理论的公理的结构,引进更复杂的方法,可能是有必要的,但目标是相同的,也就是说,根据不变性描述有意义的概念。 112

为了避免混淆,有人注意到:正是在根据表征给出这些概念时,才需要检验不变性,比如,在测量案例中的数值表征,或者,在几何案例中,用笛卡儿坐标的表征。当已知纯粹的定性关系,即用一个理论的定性的原始关系所定义的关系,例如,欧几里得几何的那些关系,那么立即随之出现的是,所定义的关系是不变的,因而是有意义的。另一方面,表征的很大的重要性是,

简化了它们完成的计算和符号以及结构的理解。这使得我们迫切需要有关于表征的不变性和意义性的明确概念，从表面上看，这可能与建立理论模型的定性结构相去甚远。

在物理学的情况下，一个理论的原始概念本身不一定是不变的。例如，如果我们在一个给定的参照系中使力学公理化，那么一个粒子的位置概念不是不变的，而是服从一种变换。在这种情况下，就要求对不变性和意义性进行一种更复杂的分析（关于这样一种分析，参见第 6 章）。然而，一般的观点是明确的：如果不伴随着对表征的不变性的研究，那么表征研究就是不完备的。我们根据测量理论的详尽分析来详述这种观点。

测量的第二个基本问题：不变性定理。对一个测量理论的表征问题的解答不完全是揭示这个理论的结构，因为从不同的测量程序给出的赋值之间，通常有形式上的不同，这是一个很平常的方法论的事实。作为一个例证，考虑下列五个陈述：

(1) 这个房间里现在的人数是 7。

(2) 司汤达(Stendhal)在 1839 年 9 月 2 日的体重是 150。

(3) 1814 年 7 月 3 日，司汤达与简·奥斯汀(Jane Austen)的体重之比是 1.42。

(4) 今天的最高温度与昨天的最高温度之比是 1.1。

(5) 今天的最高温度与昨天的最高温度差与今天的最高温度与明天的最高温度差之比将是 0.95。

只要我们作出自然假设，即，(3)用了相同的称重仪表，不管是常衡制，还是公制，(5)用了相同的温度计，不管是华氏温度计，还是摄氏温度计，那么陈述(1)、(3)和(5)的经验意义就是明确的。相反，只有当指定了测量所用的**特殊**尺度时，(2)和(4)才有一个明确的经验意义。基于这五个陈述，我们可以从形式上区分出三种测量。计数是**绝对**尺度的一个例子。一个给定的对

象集的项数是惟一确定的。根本不能任意地选择一个可利用的单位或零。相反,质量或重量的测量是比率尺度的一个例子。测量质量的任何一个经验程序都不能决定质量的单位。一个单位的选择是由一个人或一个小组在经验上作出的任意决定。当然,一旦选择了一个测量单位,比如,克或磅,就惟一地确定了宇宙中每个其他对象的质量的大小。主张这一点的另一个方面是说,质量的测量惟一地符合乘以一个正的常数。(“符合”的技术性的用法后面会变得明确起来。)距离的测量是这类测量的第二个例子。帕洛阿尔托和旧金山之间的距离与华盛顿和纽约之间的距离之比,不管测量单位是英里还是公里,都是相同的。

113

为了避免一些通常的误解,有必要详述这种主张:测量质量的经验程序没有一个能决定质量的单位。一位化学家,当用公制重量的标准系列在一个天平上测量铁盐的样品时,可能发现这种主张是令人吃惊的,因为他可能假定,公制重量的标准系列的选择,作为他的经验程序的一部分,已经把克确定为测量单位。在说服化学家相信,他的假定是不正确的时,至少有两条论证思路,应该证明是有效的。第一,指出下面这点将是适当的,把这种最终的测量转变为以磅或英两为单位的一种测量,能够表达完全相同的信息。为了以后面的这种形式表示测量,没有必要完成用天平的进一步的经验运算。但是,有人一定认为,这位化学家会很好地答复说,尽管没有必要用当前的天平完成进一步的经验运算,但从克到磅所用的换算因子,需要求助于在确定换算因子时某个标准计量局从前完成的经验运算。对他的反驳的一种分析把我们带到了论证的第二条思路,这条论证思路更深刻和更基本。首先,他求助于从前关于其他天平的经验运算,不可能支持他,因为为了证明给他用克作为测量的标准系列

进行的测量所贴的标签是正当的,也必须求助于从前完成的经验运算,即把他的重量系列的刻度指定为公制系列的那些经验运算。重要的观点是,化学家自己完成的经验运算只是确定了,铁盐样品的质量与在他的标准系列中的一个重量子集的比率。然而,产生标准系列的公司的技术员,关于导致他的标准系列刻度的经验运算,作出了相同类型的比率陈述。换言之,这位化学家的陈述是:

(6) 铁盐样品重 1.679 克。

可以被下列陈述所替代:

(7) 这种铁盐样品的质量与我的标准系列的克重之比是 1.679,而且,我的系列的制造者保证,我的克重与位于巴黎附近的国际标准计量局标准千克的铂-铱合金的质量之比是 0.001 000 0。[①]

114 温度的测量是前面提到的第三种形式独特的测量类型的一个例子。用一个温度计测量温度的经验程序,既没有确定一个单位,也没有确定一个原点。[②] 在这类测量中,任何两个间隔之比都独立于测量单位和零点。由于显而易见的理由,这种测量被称为**区间尺度**。除测量温度之外的例子还有测量时间数据、线性的位置或基数效用。[③]

运用比率和区间尺度的绝对概念,我们可以阐述精确分析测量程序的第二个基本问题:**确定由应用程序产生的测量的尺**

---

① 显然,我忽略了测量的精度问题,这个问题在科学的许多领域内具有很大的重要性。在认真讨论这个主题时,广泛地使用了第 5 章提出的概率概念。

② 很自然,我们这里排除了考虑绝对温度的测量,这种测量的零点不是任意的。

③ **基数效用**是经济学家的术语,代表效用的一种区间尺度,与效用的纯粹定序测量截然相反。

度类型。像这一节的标题一样，我们把这个问题称为一个测量理论的**不变性**问题。从数学的观点来看，确定由给定的经验运算和关系的代数所产生的测量的尺度类型，就是确定同构于一个给定的经验模型的代数的任何两个数值模型是相关的方式。例如，在质量的情况下，我们可以作出四个等价的陈述：

(8) 质量是一个比率尺度。

(9) 质量的测量惟一地符合乘以一个正数(这个数对应于单位的一种变化)。

(10) 质量的测量惟一地符合一种相似变换(这样一种变换只是乘以一个正数)。

(11) 已知测量质量的关系和运算的一个适当代数的任意模型，那么与已知模型同构的(这个代数的)任何两个数值模型，都通过一种相似变换相关联。

在(11)中"同构"的确切含义是在第3章给出的。[①] 在这一点上，惟一的目的是，为不变性问题是什么意思提供一个一般含义，(9)和(10)中的"惟一地符合"这个术语，通常是在纯数学的语境中使用的。对惟一性问题的另一种阐述是运用数学的不变性概念：*在一个已知的测量程序是不变的条件下，确定数值变换的集合*。[②] 在后面几章的详细工作中，通常没有以技术性的方式使用不变性概念，但在给定的语境中，什么定义是适当的，以及如何将所获得的惟一结果根据不变性明确地阐述出来，都将是显而易见的。

测量尺度的分类。一般的尺度概念不是我们需要以后面提

---

① 有时，当两个不同的物体具有相同的重量测量时，映射到一个数值代数是同态的，而不是同构的，第3章也说明了这种情况。

② 我不说变换*群*，因为存在着一些测量程序，对于它们来说，变换的不变集合不是一个群。

出的准确方式下定义的概念。为了达到系统地讨论惟一性问题115 的目的，我们可以把一种尺度定义为拥有相同变换特性的测量程序的一个类。已经提到了三种不同尺度的例子，即作为一种绝对尺度的计数、作为一种比率尺度的质量的测量，以及作为一种区间尺度的温度的测量。[①]

除排序之外，还有另一种类型的尺度是**任意的**。把一个数分配给一个不太精确的定序分类。莫氏硬度尺度和蒲福(Beaufort)风速标尺是第四种很弱的尺度类型的例子，当通过刮痕检验来确定相关的硬度时，按照莫氏硬度尺度，可把矿石划分为不同的等级，根据风力标尺，风的强度被分类为无风的、一级风、二级风，等等。社会科学中有许多定序尺度，下面讨论其意义。

数字有时也被用于纯粹的**分类**。例如，在一些国家，驾照上的第一个数字表明持有者居住的县。根据这样一种尺度，除了把相同的数分配给同一个县的人，把不同的数字分配给不同县的人之外，数字的分配是任意的。

最弱的尺度是，只用数字命名一个对象或一个人。这种赋值完全是任意的。选派数字和足球运动员的编号，就是这类测量的例子。这样的尺度通常被称为**定名尺度**(nominal scale)。

我们以直观的方式区分了六种类型的尺度。我们在下面几页主要限于讨论对前四种类型的分析。然而，这几页不是测量的全部尺度，我们有时也提到其他尺度。

现在，我们想根据其变换特性来描述这六种尺度中的每一种尺度。为了供将来参考，从形式上定义所提到的六种变换中的每一种变换。首先，一种绝对尺度惟一地符合**恒等变换**。

---

① 计数，等等，并不是这种绝对尺度，但在技术上，在我们的类定义的条件下，是绝对尺度的一项。这里的用法与倾向于把类看成是特性的日常用法相一致。

定义 1. 在实数集上的这个恒等变换是函数 f,使得对于每个实数 x, $f(x)=x$。

到目前为止,计数是我们的绝对尺度的惟一例子。目前,我也包括概率在内,但后面,我们将讨论通常使假设标准化的自然性,也就是,概率增加到 1。

一种比率尺度惟一地符合一个相似变换。

定义 2. 设 f 是一个实值函数,它的定义域是实数集。于是,f 是一个相似变换,当且仅当,存在着一个正实数 α,使得对于每个实数 x,

$$f(x)=\alpha x$$

质量和距离的测量已经是我们的比率尺度的两个例子。对于质量的情形来说,如果我们希望从磅变到英两,适当的相似变换是取 $\alpha=16$。如果在距离测量中,我们希望从英尺变到码,我们指定 $\alpha=\frac{1}{3}$。恒等变换是 $\alpha=1$ 的相似变换。

第三,一个区间尺度惟一地符合一个线性变换。 116

定义 3. 设 f 是一个实值函数,它的定义域是实数集。于是,f 是一个线性变换,当且仅当,存在着一个正数 α 和一个数 β,使得对于每个数 x,

$$f(x)=\alpha x+\beta$$

如果在温度测量中,我们希望从华氏度中的 x 变到摄氏度,我用了由 $\alpha=\frac{5}{9}$ 和 $\beta=-\frac{160}{9}$ 定义的线性变换。即

$$y=\frac{5}{9}(x-32)=\frac{5}{9}x-\frac{160}{9}$$

很显然,每一种相似变换都是 $\beta=0$ 的一种线性变换。

一种定序尺度惟一地符合一个单调变换。首先定义单调增加变换和单调减小变换,是方便的,而不是直接定义单调变换。

定义 4. 设 f 是一个实值函数,它的定义域是实数的某个集合。于是,f 是一个单调增加变换,当且仅当,对于在 f 定义域中的每个 x 和 y 来说,如果 $x < y$,那么 $f(x) < f(y)$。

显然,每一个线性变换都是在所有实数集上的一个单调增加变换。平方函数,即函数 f 使得

$$f(x) = x^2 \tag{1}$$

不是一个线性变换,而是在非负实数集上的单调增加。注意对于$-5<4$来说,并没有在所有实数集上的这种特性,而是

$$f(-5) = 25 > 16 = f(4)$$

重要的是意识到,一个单调增加变换不需要是通过像(1)那样的简单方程可定义的。例如,考虑集合

$$A = \{1, 3, 5, 7\}$$

而且,设 f 是在 A 上定义的函数,使得

$$f(1) = -5$$

$$f(3) = 5$$

$$f(5) = 289$$

$$f(7) = 993$$

无疑,f 是 A 上的单调递增函数,但它不满足任何一个简单方程。

定义 5. 设 f 是一个实值函数,它的定义域是实数的某个集合。于是,f 是一个单调减小变换,当且仅当,对于在 f 定义域内的每个 x 和 y 来说,如果 $x < y$,那么 $f(x) > f(y)$。

在所有实数集上的两个单调减小变换的例子是:

$$f(x) = -x$$

和

$$f(x) = -x^3 + 2$$

作为另一个例子,再考虑集合 A,而且,设在 A 上定义 f,使得 117

$$f(1) = 6$$

$$f(3) = 4$$

$$f(5) = 2$$

$$f(7) = -10$$

显然,f 在 A 上是单调减小的。

定义 6. 设 f 是一个实值函数,它的定义域是实数的某个集合。于是,f 是一个单调变换,当且仅当,f 是一个单调增加变换或一个单调减小变换。

显然,单调变换描述了定序尺度。作为定序尺度的一个重要例子,有穷弱排序的明显的不变性定理如下。

定理 1. 设 $\mathfrak{A} = (A, \leq)$ 是一个有穷弱排序。于是,$\mathfrak{A}$ 的任何两个数值表征都是通过一个单调变换相关的。

在实践中,通常方便的是,只考虑单调增加变换或单调减小变换,但是,这种限制主要是由神圣的习惯和实践促成的,而不是由经验事实的考虑激发的。

分类和定名尺度惟一地符合一个一一对应的变换,这个变换只是一个一一对应的函数。

除了基本上是由史蒂文斯(Stevens, 1947)提出的这六种经典的尺度类型之外,还有其他一些尺度类型,将在后面讨论。一个例子是我们称之为**超定序尺度**(hyperordinal scales)的那些程序的类。它们通过保留一阶差分和我们为此贴上**超单调的**

标签的变换来描述。这里,我们可大致地把它们描述为是只缺少区间尺度的经验结构。测量感觉强度或效用的不同方法提供了某些例子。

定义 7. **设 f 是一个实值函数,它的定义域是实数的某个集合。于是,f 是一个超单调变换,当且仅当,f 是一个单调变换,而且对于在 f 定义域内的每个 x、y、u 和 v 来说,如果**

$$|x-y|<|u-v|$$

那么

$$|f(x)-f(y)|<|f(u)-f(v)|$$

很自然,每一个线性变换都是一个超单调增加变换,但是,反之则不然。例如,考虑

$$A=\{1,2,3,8\}$$

和函数 f,使得

$$f(1)=1$$

$$f(2)=2$$

$$f(4)=5$$

$$f(8)=15$$

118 显然,f 是超单调增加的,但在 A 上不是线性的。

严格地说来,存在着一个由数值变换的各种群来描述的不可数的无穷多的尺度,但是它们中的大多数没有任何真正的经验意义。此外,把测量概念推广到有序结构,比如,晶体和不能以自然方式用数字表示的部分排序,是有可能的。我自己的观点是,这种推广是不受欢迎的,因为在斯科特和苏佩斯(1958)的语言中,它把测量理论的类等同于有穷关系结构的类,而且有穷关系结构的一般理论远远超越了任何一种直

观的测量概念。[①]

把这一节的分类分析总结为表2。

表2 尺度和测量的分类

尺 度	变 换 特 性	事 例
绝对的	恒等	计数,相对频率
比率的	相似 (乘以一个正数)	质量、距离、电流、电压
区间的	线性的(乘以一个正数和加上一个任意数)	温度、势能、线性的位置、基数效用
超定序的	超单调的(保留一阶差分的顺序)	音调、响度、效用
定序的	单调或单调增加	莫氏尺度、蒲福风级尺度、定性的偏爱
定名的	一一对应	电话号码、社会安全号

现在,我转向与3.4节的四个测量表征定理相对应的不变性定理。

定理2. 设 $\mathfrak{A}=(A, \mathcal{F}, \succeq)$ 是一个有穷的等空间的可扩展的结构。于是,$\mathfrak{A}$ 的一个数值表征,在表2的相似变换群的条件下,是不变的。

这个基本定理的证明留给读者。

定理3. 设 $\mathfrak{A}=(A, \succeq)$ 是一个有穷的等空间的差结构。于是,$\mathfrak{A}$ 的一个数值表征,在表2的线性变换群的条件下,是不 119

① 也许正如显而易见的那样,一个有穷的关系结构是它的定义域是有穷集的一个关系结构。已知这种有穷性,能够以或多或少是任意的方式来建构某种数值表征,因而,这没有太大的(如果有一点的话)科学意义。只有那些反映实际测量程序(也许在某种程度上是理想化的)的数值表征,才有科学意义。

变的。①

定理 4. 设 $\mathfrak{A}=(A, \geq, B)$ 是一个有穷的对分结构。于是，$\mathfrak{A}$ 的一个数值表征，在表 2 的线性变换群的条件下，是不变的。②

---

① 定理 3 的证明。为了证明定理 3 的数值函数 φ 惟一地符合一个线性变换，我们定义，对于 A 中的每个 a 来说，两个函数 h_1 和 h_2 为：

$$h_1(a)=\frac{\varphi_1(a)-\varphi_1(c^*)}{\varphi_1(c^*)-\varphi_1(c^{**})}$$

$$h_2(a)=\frac{\varphi_2(a)-\varphi_2(c^*)}{\varphi_2(c^*)-\varphi_2(c^{**})}$$

其中，φ_1 和 φ_2 是满足表征结构的两个函数，c^* 是在排序$\geq$条件下 A 的第一个元素，c^{**} 是第二个元素。我们能容易表明，h_1 是 φ_1 的一个线性变换，h_2 是 φ_2 的一种线性变换，还有，h_1 等同于 h_2。然后，很容易证明，φ_1 是 φ_2 的一个线性变换，即存在数 α、β，且 $\alpha>0$，使得对于 A 中的每个 a，

$$\varphi_1(a)=\alpha\varphi_2(a)+\beta$$

严格地说来，在排序$\geq$的条件下，c^* 不一定是这种第一个元素，c^{**} 也不一定是这种第二个元素，因为在定义域 A 中，可能有其他项在这个排序中等价于 c^*，也有其他项等价于 c^{**}。使这一点变得更精确的最简单的技术方式是，引入关于这种排序的等价类，正如第 3 章的一个脚注中证明定理 3 时所做的那样。

② 定理 4 的证明。像证明定理 3 的情形那样，为了证明定理 4 的数值函数 φ 惟一地符合一个线性变换，我们假定，我们有两个函数 φ_1 和 φ_2 都满足(i)和(ii)。于是，恰好像在那个证明中那样，我们定义 h_1 和 h_2。通过这种定义方式，显然，h_1 是 φ_1 的一个线性变换，h_2 是 φ_2 的一个线性变换。我们通过归纳论证完成的证明表明，$h_1=h_2$（因此，φ_2 是 φ_1 的一个线性变换）。

这种归纳是通过$\succ$排序关于 A 的元素，这个序列的第一个元素是 a_1。

现在，根据定义

$$h_1(a_1)=h_2(a_1)=0$$

假定现在对于满足 $m\leqslant n$ 的 a_m，

$$h_1(a_m)=h_2(a_m)$$

我们证明

$$h_1(a_{n+1})=h_2(a_{n+1})$$

现在我们立即知道，$a_{n-1}Ja_n$和a_nJa_{n+1}，由此，根据公理 5

$$B(a_{n-1}, a_n, a_{n+1})$$

（转下页）

值得注意的是，下一个不变性定理（定理 5）拥有一种自然几何的解释。如果我们考虑成对地映射到笛卡儿平面的函数 φ_1 和 φ_2，那么，惟一性定理指出，在标准几何意义上，尺度的任何变化一定在各个方向上都是均匀的，但是，沿着不同的轴平移不同的距离，可能会使原点平移。

定理 5. 设 $(A_1, A_2, \geq)$ 是一个非平凡的有穷的等空间的相加联合结构。于是，像在定理 3.4.4 中那样，一个数值表征 (φ_1, φ_2) 是不变的，取决于 (φ_1, φ_2) 的线性变换的对 (f_1, f_2) 的群，但是，受到的限制是，f_1 和 f_2 具有一个共同的乘数，即存在着数 α、β 和 γ，使得 120

$$f_1 \circ \varphi_1 = \alpha\varphi_1 + \beta \text{ 和 } f_2 \circ \varphi_2 = \alpha\varphi_2 + \gamma \text{①}$$

（接上页）因此，通过假设

$$2\varphi_i(a_n) = \varphi_i(a_{n-1}) + \varphi_i(a_{n+1})$$

由此，

$$\varphi_i(a_{n+1}) = 2\varphi_i(a_n) - \varphi_i(a_{n-1})$$

其中，$i = 1, 2$。现在，既然 h_i 是 φ_i 的一种线性变换，所以，接下来我们也有

$$h_i(a_{n+1}) = 2h_i(a_n) - h_i(a_{n-1})$$

但是，通过归纳假设，对于 h_1 和 h_2 来说，最后这个方程的右边是相同的，这样，我们断定

$$h_1(a_{n+1}) = h_2(a_{n+1})$$

① 定理 5 的证明。为了证明 φ_1 和 φ_2 的不变结果，我们可以仿照定理 3 的情形。我们定义四个函数：

$$g(a) = \frac{\varphi_1(a) - \varphi_1(c^*)}{\varphi_1(c^*) - \varphi_1(c^{**})} \qquad g'(a) = \frac{\varphi_1'(a) - \varphi_1'(c^*)}{\varphi_1'(c^*) - \varphi_1'(c^{**})}$$

$$h(p) = \frac{\varphi_2(p) - \varphi_2(r^*)}{\varphi_2(r^*) - \varphi_2(r^{**})} \qquad h'(p) = \frac{\varphi_2'(p) - \varphi_2'(r^*)}{\varphi_2'(r^*) - \varphi_2'(r^{**})}$$

其中，c^* 是在 A_1 上进行排序 $\geq$ 条件下 A_1（原文是 A 有误，应该是 A_1。——译者）的第一个元素，c^{**} 是第二个元素；r^* 是 A_2 的第一个元素，r^{**} 是第二个元素。像以前那样，显然，g 是 φ_1 的一个线性变换，g' 是 φ_1' 的一个线性变换，h 是 φ_2 的 （转下页）

4.4 物理学理论的基本方程为什么不是不变的?

考虑到在物理学中不变性概念是重要的,甚至具有深刻的物理意义,通常在数学中也是如此,那么,自然要问的问题是,理论为什么很少是以不变的形式写成的。

我的回答将集中于物理学,也稍微更普遍地说集中于测量理论。在物理学中一个明显的并在实践中重要的观点是,相对于实验室的固定框架完成并记录测量,而不是以经典的或相对

(接上页)一个线性变换,h'是φ_2'的一个线性变换。第二,我们通过类似于定理4用过的归纳论证,能够表明,$g = g'$和$h = h'$。因此,我们得到,存在着数α、α'、β和γ,具有α, $\alpha'>0$,使得对于A_1中的每个a和A_2中的每个p,

(iii) $\varphi_1'(a) = \alpha\varphi_1(a) + \beta$和$\varphi_2'(p) = \alpha'\varphi_2(p) + \gamma$。

这仍然表明,当A_1和A_2每个都在序列中至少有两个不等价的元素时,$\alpha = \alpha'$。我们可以不失一般性地假定,$a > b$和$p > q$。那么我们从$(a, q) \approx (b, p)$有

$$\varphi_1'(a) - \varphi_1'(b) = \varphi_2'(p) - \varphi_2'(q)$$

这样,通过(iii)

$$\frac{\alpha\varphi_1(a) - \alpha\varphi_1(b)}{\alpha'\varphi_2(p) - \alpha'\varphi_2(q)} = 1$$

所以

$$\frac{\alpha}{\alpha'}\left(\frac{\varphi_1(a) - \varphi_1(b)}{\varphi_2(p) - \varphi_2(q)}\right) = 1$$

但是,通过假设

$$\varphi_1(a) - \varphi_1(b) = \varphi_2(p) - \varphi_2(q)$$

由此,

$$\frac{\alpha}{\alpha'} = 1$$

即

$$\alpha = \alpha'$$

这就完成了这个证明。

论的不变方式记录测量,更简单和更方便。

此外,如果我们考察测量单位的使用,这种观点还要更加明显。只分析长度和质量的测量就足够了。在这两种情况下,测量单位的选择都是任意的,即不是不变的。不变的是距离或质量的比率。在古希腊的数学中,能够看到不断地使用比率的不得体的惯用语,而不是约定的单位标准的简写。事实上,基本的计算显示了单位的优势。假定我们需要比较一千个城市与机场之间的距离。如果只能记录不变的比率,那么,与一千个物理量,即一千个有共同单位的数字(比如,米或公里)正相反,将会记录下作为纯粹数字的大约五十万个比率。 121

热衷于以希腊的风格写比率持续了很长时间。牛顿的《自然哲学的数学原理》(1687/1946,第 13 页)是一个极好的例子。这里是牛顿对他的运动第二定律的经典几何学的阐述:"运动的变化正比于所施加的运动的力;而且,沿着作用力的直线方向发生变化。"有人通常主张,并且相当敏锐地说,当在物理学中引入更可理解的和更有效的阐述与计算方法时,从表达的希腊几何风格转变到代数方程和微分方程,具有很大的重要性。

除了相当特殊的比率情形之外,在标准几何的综合的数学表述和分析的数学表述中,同样的区别也是明显的。欧几里得几何的一种综合的和不变的数学表述,很容易只用介中性和全等术语的定性关系加以阐明。相比之下,分析表征是相对于任意选择的坐标系的,因而,完全是可变的。通常,证明一个不变性定理是表明,各种不同的适当分析的表征如何,即通过哪些变换关联起来。事实上,这些变换构成了已经提到的欧几里得的运动群的分析形式。

在物理学中,普遍承认分析表征的优势,特别是在矢量空间方面的分析表征优势。人们认为,推荐返回到 18 世纪之前所用

的综合几何的数学表述的任何一位物理学家，都是极其特殊的和古怪的。

置身于这种语境中，人们能够立即明白，在作为综合的数学表述方面的不变性和作为分析的数学表述中的有效计算之间，存在着一种无法避免的矛盾。事实上，对方便计算和在实践中简化计算的"正确"坐标系的选择，被看成对物理学的教学很有意义。①

122 **超越对称性**。从我所说的来看，可能会令人误导地推出，在选择方便的实验室的坐标系时，物理学家忽略了考虑不变性。在实践中，物理学家通过引入和运用协变性概念来坚持不变性。在引入这个概念之前，超越熟悉的几何学中的对称性，考察一下自然的概括，是有用的。

欧几里得空间的最一般的自同构是由旋转、偏转和平移构成的变换。我们已经在 4.1 节中考虑了旋转和偏转。像圆、正方形和等边三角形之类的常见图形，在适当的旋转和偏转群的条件下，是不变的——对于圆来说，是一个无穷群；对于正方形和等边三角形来说，是一个有穷群。

另一方面，这些常见的图形，在包括平移在内的自同构的条

---

① 在经典物理学中，通常归因于西亚斯(Siacci，1878)的一个漂亮的运动学的例子是，把一个运动质点的解，限于使用径矢的一个平面，在每个时间 t，从这个平面的一个确定的原点到质点的位置。这个矢量用两个分量表示，不同于平常的(x, y)坐标，而是一个沿着切线，另一个沿着这个矢量的法线，即垂直于切线。现在，考虑一个质点，它的加速度的法线分量和切线分量不随时间变化。因此，我们有了沿着两个运动轴的常数加速度，而且，在这种表征中，很容易求解运动方程来表明，它在每一时刻的位置 s 通过下式给出：

$$s = ae^{b\varphi}$$

其中，φ 是切线与所确定的线之间的夹角。这个例子的细节能够在惠特克(Whittaker，1904/1937，pp. 21 - 22)的著作中找到。

件下,没有一个是不变的。事实上,所有的点都是可变的,所以,惟一不变性是,转移到自身的整个空间的不变性。但是,平移不仅在纯几何学中是重要的,而且在应用几何学中,特别是在物理学中,也是重要的。在物理学中,经典物理学的伽利略变换和狭义相对论的洛伦兹变换发挥着重要的作用,这两种变换都包括平移在内。

在最后一段,从以像介中性和全等之类的原始概念为基础的欧几里得几何的定性的自同构,转向伽利略和洛伦兹变换群的数值坐标参照系,作出了一个不引人注意的转移。为了对物理学家的一个理论的协变性概念给出明确的说明,需要这样的参照系。

这样的协变性的典型例子是速度和加速度,不管是在伽利略变换的条件下,还是在洛伦兹变换的条件下,从一个参照系到另一个参照系,速度与加速度都不是不变的。此外,一个粒子的速度和加速度的矢量方向,在一般情况下,将从一个参照系变化到另一个参照系。(加速度的标量值是不变的。)

协变性。物理学的定律就是根据这样的协变性写成的。如果不打算进行最一般的数学表述的话,那么基本思想可表达如下:设,Q_1, …, Q_n 是时空坐标函数的参量,例如,其中的某些参量 Q_i 是其他参量派生的。那么,一般情况下,当我们从一个坐标系转到另一个坐标系时,Q_1', …, Q_n' 将是协变的,而不是不变的,因此,它们的数学形式在新的坐标系中是不同的。但是,在新的坐标系中,包括这些量的任何一个物理定律,比如说,

$$F(Q_1, \cdots, Q_n) = 0 \tag{1}$$

一定有相同的形式:

$$F(Q_1', \cdots, Q_n') = 0 \tag{2}$$

这种相同形式的要求,是重要的不变的要求。用通常的物理学家的语言来说,方程(1)和(2)也被称为是协变式。因此,**协变性**这个术语,既适用于参量,也适用于方程。我省略了对这些熟悉思想的明显的形式陈述。

这里是来自经典力学的一个例子。考虑两个粒子在碰撞前后的动量守恒,v_i 是碰撞前的速度,w_i 是碰撞后的速度,m_i 是每个粒子 $i=1, 2$ 的质量。根据形式(1),这个定律看起来似应如此:

123

$$v_1 m_1 + v_2 m_2 - (w_1 m_1 + w_2 m_2) = 0$$

当然,变换后的形式将是:

$$v_1' m_1 + v_2' m_2 - (w_1' m_1 + w_2' m_2) = 0$$

但一般情况下,速度 v_i 和 w_i 将是协变的,而不是不变的。质量 m_1 和 m_2 当然是不变的。

最后,我提到的是,在复杂问题中,证明一个物理学定律是协变的,即它的形式的不变性,可能至少是相当乏味的。

4.5 在各态历经理论中作为完全不变的熵[①]

现在,我转向一个扩展的例子,它是在数学与物理学中都具有结果的不变性的最漂亮的案例。我们需要为这种讨论提供某些方法,这预示了第 5 章第一节的详细进展。

---

① 这一节在不失连续性的条件下可以忽略掉。关于这里概述的结果的一个好的详细参考文献和在各态历经理论中对不变性的一个更一般的考察,将会在奥恩斯坦(Ornstein)和外斯(Weiss)的著作中(1991)找到。这一节的概率概念的用法超出了第 5 章引入的概率的范围,第 5 章集中于概率的基础,而不是概率的应用。

让我们首先从一个标准的概率空间(Ω, $\mathfrak{I}$, P)开始,不用说,在这里,$\mathfrak{I}$ 是 Ω 的子集的一个 σ -代数,P 是一个在 $\mathfrak{I}$ 上的 σ-相加的概率测度。(参见 5.1 节的定义 2)我们现在考虑从 Ω 到 Ω 的一种映射 T。我们说,T 是**可测的**,当且仅当,只要 $A \in \mathfrak{I}$, 那么 $T^{-1}A = \{\omega: T\omega \in A\} \in \mathfrak{I}$, 乃至更重要的是,$T$ 是**保测的**,当且仅当, $P(T^{-1}A) = P(A)$。映射 T 如果拥有下列三个条件,那么就是**可逆的**:(i) T is $1-1$,(ii) $T\Omega=\Omega$,以及(iii) 如果 $A \in \mathfrak{I}$, 那么 $TA = \{T\omega: \omega \in A\} \in \mathfrak{I}$。在这种应用中,我们感兴趣的是,$\Omega$ 中每个 ω 都是一个双重的无穷序列,而且,T 是**右移的**,使得,如果对于所有的 n, $\omega_n = \omega'_{n+1}$, 那么 $T(\omega) = \omega'$。在直观上,这种特性对应于过程的平稳性——时间的推移不影响这个过程的概率定律,然后,我们能够用 T 描述 Ω 中的轨道或样本路径。

我现在为各态历经理论引入熵的核心概念。为了保持数学概念和符号的简单性,我限于离散时间过程,事实上,也限于含有有限数目的状态的过程。因此,对于这样的过程来说,我们有一个简单的熵定义。

首先,一个具有(离散的)概率密度 $p(x)$的随机变量 X 的熵通过下列等式来定义:

$$H(\boldsymbol{X}) = -\sum p_i \log p_i$$

在类似的思路中,对于一个随机过程 $\boldsymbol{\chi} = \{\boldsymbol{X}_n: 0 < n < \infty\}$,

$$H(\boldsymbol{\chi}) = \lim_{n \to \infty} \frac{1}{n} H(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n) \qquad (1)$$

注意我们所做的是,把这个过程的熵定义为联合分布的熵的极限。为了阐明这里的思想,我们能够马上写出具有一个联合离散密度分布 $p(\boldsymbol{X}, \boldsymbol{Y})$的一对离散随机变量的联合熵 $H(\boldsymbol{X}, \boldsymbol{Y})$ 124

的明显表达式。

$$H(\boldsymbol{X}, \boldsymbol{Y}) = -\sum_{x}\sum_{y} p(x, y)\log p(x, y)$$

这也能够被表示为

$$H(\boldsymbol{X}, \boldsymbol{Y}) = -E\log p(\boldsymbol{X}, \boldsymbol{Y})$$

其中,E 代表期望值。当然注意,我关于(1)的必要条件有点随意。重要的是需要有这种限制。因此,当 n 趋于无穷大时,如果不存在有穷联合分布的熵的限制,那么随机过程的熵就是不可定义的。

对于一个伯努利(Bernoulli)过程来说,即每个试验都是同一的与独立的分布随机变量(i. i. d.)的一个过程,我们有一个特别简单的熵表示。像下列方程所表示的那样,它只是一个随机变量的熵:

$$H(\boldsymbol{\chi}) = \lim_{n\to\infty}\frac{1}{n}H(\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n)$$

$$= \frac{nH(\boldsymbol{X}_1)}{n} = H(\boldsymbol{X}_1) = -\sum p_i \log p_i$$

对于马尔可夫链的情形来说,我们只需要一个稍微更复杂的定义,此外,当然根据规定,存在着这种限制。

$$H(\boldsymbol{\chi}) = \lim_{n\to\infty}H(\boldsymbol{X}_n \mid \boldsymbol{X}_{n-1}, \cdots, \boldsymbol{X}_1)$$

$$= H(\boldsymbol{X}_2 \mid \boldsymbol{X}_1)$$

$$= -\sum_i p_i \sum_j p_{ij}\log p_{ij}$$

现在,由于例子很简单,所以,我们在没有作出详尽的技术说明的前提下,给出了各态历经的定义。设$(\Omega, \mathcal{F}, P, T)$是具有保测变换 T 的一个概率空间。如果对于任何一个可测集 A,

$\mu(TA) = \mu(A)$，那么 T 就是**平稳的**，而且，如果对于每个集合 A，使得 $TA = A$，即，A 的测度要么是 0，要么是 1，那么变换 T 被称为**各态历经的**。我们通过根据 T^n 定义随机变量 $\boldsymbol{X}_n$，获得了一个随机过程，即对于 Ω 中的每个 ω，$\boldsymbol{X}_n(\omega) = \boldsymbol{X}(T^n\omega)$。很容易看出，上面定义的伯努利过程是平稳的，也是各态历经的。对一个随机过程的平稳性的一个更直观的描述是，这个过程的随机变量的有穷子集的每一种联合分布，在时间平移的条件下，都是不变的。因此，对于每个 k，

$$p(x_{t_1+k}, x_{t_2+k}, \cdots, x_{t_m+k}) = p(x_{t_1}, x_{t_2}, \cdots, x_{t_m})$$

也许，一个非各态历经过程的最简单的例子是在硬币的正反面中的马尔可夫链：

	h	t
h	1	0
t	0	1

已知初始概率是 $p_1(h) = p_1(t) = \frac{1}{2}$，在每次投掷硬币的试验中，由具有正面的样本路径组成的事件 A 有概率 $\frac{1}{2}$，即 $p(A) = \frac{1}{2}$，但是，$T(A) = A$，因此，这个过程不是各态历经的。在直观上，从这个例子已经搞清楚，我们从各态历经过程所预期的是：消除所有依赖于初始态的混合态。 125

另一方面，当这个矩阵是稍微不同的，而且，在第一次试验中出现正面的概率现在是 1 时，我们有一个各态历经过程，但熵是 0。

	h	t
h	0	1
t	1	0

$p(h_1) = 1$

从一个可能序列 $hththt$ ... 足以明白,所有的测量或结果都是可预言的,因此,熵一定是 0。注意这个过程是什么。它不是一个伯努利过程,而是一个有两种状态的非常特别确定的马尔可夫链。与这个例子相关,我提到,文献中对各态历经的马尔可夫链的标准定义没有过分地要求平稳性。通常所用的定义是,一个马尔可夫链是各态历经的,如果它有独立于初始分布的态的惟一的渐近概率分布。因此,关键的是初始分布的独立性,而不是平稳性。当然,在这里,显而易见的是,这个过程必须逐渐平稳地成为各态历经的。

当一个随机过程有正熵时,我们能够作出某些进一步重要的区分;不是所有的测量都是可预言的,只是某些测量可能是可预言的。例如,我们能够确定这个随机过程的一个因数,即把随机过程限于事件的子代数。在这里所用的概念中,一个因数总是恰好是事件的这样一个子代数。一个 K 过程(K 属于柯尔莫哥洛夫过程)的概念是一个随机过程的概念:这个随机过程没有 0 熵的因数,即没有确定的因数。伯努利过程是 K 过程,但也有许多其他过程。

关于各态历经过程的一个重要事实是基本的各态历经定理。为了阐明这个定理,我们需要一个集合 A 的指标函数的概念。对于任何一个事件 A 来说,设 I_A 是 A 的指标函数。即对于 $\omega \in A$,

$$I_A(\omega) = \begin{cases} 1 & \text{如果 } \omega \in A \\ 0 & \text{如果 } \omega \notin A \end{cases}$$

基本的各态历经定理的含义是,时间平均值等于空间或系综平均值,也就是说,我们通过考虑一个样本路径 ω 的转移或变换,恰好得到了同样的东西,而不是考察具有事件 A 所表达

的一个特性的所有 ω 的一个时间截面。对于除测度 0 的集合之外的所有 ω 来说，这个定理都成立。

定理 1. 对于 Ω 中的几乎任何一个 ω，

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}I_A(T^k\omega)=P(A)$$

各态历经过程的同构。如果我们能够把一组事件映射到另一组事件，而保持概率不变，那么我们就把两个平稳的随机过程定义为（在测量定理的意义上）同构的。这一点的直观意义是，它们的结构的不确定性是一样的。

更技术性地说，我们以下列方式定义同构。已知两个概率空间 $(\Omega, \mathfrak{I}, P)$ 和 $(\Omega', \mathfrak{I}', P')$，以及分别在 Ω 和 Ω' 上的两个保测变换 T 和 T'，那么，我们说，$(\Omega, \mathfrak{I}, P, T)$ 在测量定理的意义上同构于 $(\Omega', \mathfrak{I}', P', T')$，当且仅当，存在着一个函数 φ: $\Omega_0 \rightarrow \Omega_0'$，这里，$\Omega_n \in \mathfrak{I}$, $\Omega_n' \in \mathfrak{I}'$, $P(\Omega_0) = P(\Omega_0') = 1$，满足下列条件： 126

(i) φ 是一一对应的；

(ii) 如果 $A \subset \Omega_0$ 和 $A' = \varphi A$，那么，$A \in \mathfrak{I}$，当且仅当，$A' \in \mathfrak{I}'$，而且，如果，$A \in \mathfrak{I}$

$$P(A) = P'(A')；$$

(iii) $T\Omega_0 \subseteq \Omega_0$ 和 $T'\Omega_0' \subseteq \Omega_0'$；

(iv) 对于 Ω_0 中的任意 ω，

$$\varphi(T\omega) = T'\varphi(\omega)$$

任何一个事件 A 到相同概率的一个事件 A' 的映射，即条件(ii)在概念上是很关键的，在(iv)中所表示的交换性也是这样。

很重要的一个问题一直是理解随机过程如何通过它们的同构关联起来。直到 20 世纪 50 年代中期，一个有争议的问题是，

按照上面给出的定义，下列两个伯努利过程是否是同构的：一个过程有两个态，每个态的概率都是二分之一，$B\left(\frac{1}{2}, \frac{1}{2}\right)$，另一个过程有三个态，每个态的概率都是三分之一，$B\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$。这个问题很容易陈述，但很难解决。最后，下列定理是由柯尔莫哥洛夫和西奈(Sinai)证明的。

定理 2. (Kolmogorov 1958,1959;Sinai 1959)如果两个伯努利过程是同构的，那么，它们的熵是相同的。

既然 $B\left(\frac{1}{2}, \frac{1}{2}\right)$和 $B\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$有不同的熵，所以，与定理 2 相反，它们不是同构的。直到十二年之后，奥恩斯坦才证明了这个逆命题。

定理 3. (Ornstein 1970)如果两个伯努利过程有相同的熵，那么它们是同构的。

根据这两个定理，扩展到其他过程，特别是马尔可夫过程，是相对简单明了的。因此，容易证明下列定理。

定理 4. 任何两个不可简化的、平稳的、有穷态的离散的马尔可夫过程是同构的，当且仅当，它们有相同的周期和相同的熵。

从这个定理得到了令人惊奇的和重要的不变性结果。对于一个非周期的、各态历经的马尔可夫过程的测量理论的同构而言，熵是完全不变的。成为完全不变的也意味着，具有相同周期的两个马尔可夫过程是同构的，当且仅当，它们的熵是相同的。这个结果令人惊奇和重要之处是，一个马尔可夫过程的概率结构是大的和复杂的，另一方面，这个熵是一个实数。对于下列有意义的数学结构来说，很难想到一个不变性的结果：在这些结

构中,结构的复杂度和完全不变的随意陈述之间的差异是如此之大。当然,在概率论中,作为一个重要的不变性结果,是很难相匹配的。根据完全不变的本性,在可比较简单性的物理学中,这种不变性结果也是难以想到的。

另一方面,重要的是注意到,甚至一阶各态历经的马尔可夫过程和一个伯努利过程可能有相同的熵之比,因而,在测量理论的意义上是同构的,但它们也绝不是同类过程。例如,我们能够通过很直接的统计检验表明,一个足够长的已知的样本路径是来自伯努利过程,还是来自一阶的马尔可夫过程。例如,存在着把这两个过程区别开来的一种简单的卡方检验。在这个过程中,这是检验一阶与零阶依赖性。 127

但是,这种情境在不变性理论中是很常见的。我们作为惯性系的观察者,肯定能够在我们的惯性参照系中是静止的一个粒子与以大于零的恒定速度运动的一个粒子之间作出区分。对于两个不同惯性系的观察者(伽利略的或洛伦兹的)来说,他们所观察到的在现象学意义上可能是完全不同的。这种阐述问题的方式可能似乎不同于几何的或各态历经的情形中所用的进路,但这只在表面上是如此。在更深的和更一般的层次上,这里概述的所有的不变性情形都有的一个共同主题是:辨别恒定的或不变的某些特殊特性。

5.
概 率 的 表 征

129　　关于概率基础的当代主要思想流派都没有对第1节考虑的任何一个公理、定义或定理提出质疑,仅注意到极少数的几个例外。从第1节介绍的概念观点来看,我们把关于概率基础的长期争论阐述为关于下列问题的争论:把定义2描述的概率论的哪些模型看成是真正的概率空间。换句话说,这种争论的核心是围绕如何把定义2的公理扩展到只包括"真正的"概率空间展开的。在这一节和接下来的几节,我们将把关于概率本性的许多不同立场划分为六种最重要的意见类。第2节致力于经典定义,拉普拉斯在19世纪初曾对此进行过相对明确的阐述。下一节致力于概率的相对频率的观点,19世纪末,约翰·维恩(John Venn)首先以相当明确的方式阐述了这种观点,后来,它提到了理查德·冯·米泽斯(Richard von Mises)等人的研究与推广。第4节聚焦于更新近的有穷序列中的随机理论和作为柯尔莫哥洛夫定义的这些序列的复杂度的测量。第5节致力于概率的逻辑理论,特别强调了杰弗里斯(Jeffreys)的工作和卡尔纳普等人提出的确证理论。

倾向性理论,尤其是以具体的物理学或心理学

通常没有注意到的定性的不变关系。第7

与德·菲内蒂(de Finetti)、古德(I. J.

Good)、林德利(Lindley)、萨维奇(L. J. Savage)的工作特别相关。第 8 节,即最后一节,概述了关于概率本性的实用性观点,这种观点不完全支持前面几节分析的六种主要立场中的任何一种立场。

关于概率基础的统计文献比其哲学文献要多,这些文献与统计推理基础或某些哲学家所称的归纳推理密切相关。因为关于这两个相关的基本论题的文献现在是如此之多,所以,我在这一章只限于讨论概率的基础,在概率的本性与统计推理的特殊形式纠缠不清的地方,只对统计推理进行一些粗略的评论。当前的这一章已经很长,可还是远远不足以表达概率的不同表征或解释的许多重要方面。需要有同样长的一章,才能对统计基础的可替代 130 进路作出比较分析。

5.1 形式理论

我们对概率的公理化是通过定义集合论谓词“是一个概率空间”进行的。提出这种观点的智力源头是柯尔莫哥洛夫的经典著作(1933/1950)。本节的主要目的是提出概率理论的公理与概念,并作为本章的其余部分研究概率基础的背景。熟悉标准形式发展(像在现代概率课上所讲的那样)的读者,应该跳过本节。所有关于概率的哲学讨论几乎都置于本章的最后。此外,某种更不标准的形式理论将会出现在关于量子力学的隐变量理论的 7.2 节、关于随机过程的可逆性概念的 7.3 节以及关于随机学习模型的 8.3 节和 8.4 节。

原始概念。公理以三个原始概念为基础:可能结果的一个非空集 Ω、表征可能事件的一个子集族 $\mathfrak{J}$,[1]以及在 $\mathfrak{J}$ 上的一个

---

① 在下文中,可能事件被简单地称为“事件”。

实值函数 P,即它的值域是实数集的一个函数;至于 $E \in \mathfrak{I}$, $P(E)$ 被解释为 E 的概率。这三个概念可以通过一个简单的例子加以说明。设 Ω 是把一枚硬币掷三次的所有可能结果的集合。

$$\Omega = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$$

这里,hhh 是连续三次获得正面的事件,以此类推。设 $\mathfrak{I}$ 是 Ω 的所有的子集族。至少得到两次正面的事件是集合

$$A = \{hhh, hht, hth, thh\}$$

恰好得到两次正面的事件是集合

$$B = \{hht, hth, thh\}$$

等等。这是重要之点。把这枚硬币掷三次的结果能发生的任何一个事件,都可以通过 Ω 的一个子集来表示。这个子集含有像刚才的那些可能结果的元素,其中任何一个可能发生的结果都包含着这个事件的发生。因此,如果发生四种可能结果中的任何一种结果,那么就发生了至少得到两次正面的事件。

作为集合的事件的这种表征,是把概率理论应用于经验数据的一个更重要的方面。确定集合所表示的事件,就是作出关于实验数据的标准形式的主要决定。正如我们在把一枚硬币掷三次的简单情形中能够看到的那样,我们表示实验结果的方式,忽略了实际掷硬币的大多数物理细节。例如,我们没有记录硬币静止下来的时间,也没有记录两次投掷之间的时间间隔。我们没有描述用来投掷硬币的机制。我们只对实际发生事件的一种很简单的而且事实上是很纯朴的抽象感兴趣。

131 现在,在关于概率理论的教科书中,作为所有可能结果集合的子集的事件语言和事件表征,都是标准的。也许,惟一主要的例外会在哲学家和逻辑学家提供的关于概率的逻辑理论的文献中找到;但正如我们在后面会看到的那样,在决不影响该理论的

重要内容的前提下,提供它的集合论版本,是轻而易举的。然而,大多数人总是不接受这种事件语言。这种语言的早期困难是由于它缺乏明确定义的形式基础,就像现在用到的标准的集合论解释所提供的那样。在描述作为所有可能结果的子集的事件之前,或者,正如我们通常指出的那样,在描述作为样本空间或概率空间的子集的事件之前,一个事件的逻辑结构没有得到很好的定义,而且,总是搞不清楚把什么算作是一个事件。① 在凯恩斯(Keynes)关于概率的一篇文章的下面一段话中(1921, p. 5)对一个事件的发生概念的含糊性提出了异议。

> “事件”这个术语迄今为止在这门学科的用语中一直占有如此重要的地位,我一概不用这个术语。论述概率的作者一般情况下研究他们所称的“事件”的“发生”。在他们首先研究的问题中,这并没有更多地违反常见的用法。但是,现在所用的这些表达,在某种程度上,是模糊的和不明确的;于是,用讨论命题的真理和概率取代了讨论事件的发生和概率,并不只是一种言语的改进。

值得注意的是,在凯恩斯的异议出现十多年之后,事件概念才在柯尔莫哥洛夫的经典论文中得到了澄清。凯恩斯自己的概率观中的许多困难是由命题概念的含糊性引起的。这个概念的基本内涵特征使其很难在概率理论中作为一个基本概念来使用,如果随后应用了标准的数学方法的话。我在第5节讨论概率的逻辑理论时再一次考虑这一点。

让我们回到掷硬币的例子。如果硬币是公平的,那么对于

---

① 我没有在“样本空间”和“概率空间”之间作出形式的区分,但很自然是,一个实验的样本空间通常是有穷的,而一个理论的概率空间通常是无穷的。接下来的集合论定义(定义2)只涉及“概率空间”。

Ω 中的每个 ω,

$$P\{\omega\} = \frac{1}{8}$$

通过合计这个子集中的元素的总数,就可以简单地获得任何其他事件的概率。因此,在掷三次硬币时至少得到两次正面的概率是 $P(A)$,而且,

$$\begin{aligned} P(A) &= \{hhh,\ hht,\ hth,\ thh\} \\ &= P\{hhh\} + P\{hht\} + P\{hth\} + P\{thh\} \\ &= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2} \end{aligned}$$

刚才举例说明的相加性是假定为一个公理的概率的基本特性之

132 一:两个相互排斥的事件的任何一个概率都等于各个概率之和,也就是说,如果 $A \cap B = \varnothing$, 那么

$$P(A \cup B) = P(A) + P(B)$$

事件的语言。既然我在本书的其余部分从头到尾都把概率语言应用于解释为事件的集合,所以,用一张表把集合论的符号与概率术语联系起来是有用的。首先,断言 $A \in \mathfrak{I}$,对应于断言,A 是一个事件。乍一看,子集族 $\mathfrak{I}$ 的概念似乎是多余的。把样本空间 Ω 的所有可能的子集看作事件,似乎是合理的。当 Ω 是一个有穷集时,这是一条合理而自然的进路;但是,对于在概率理论的许多应用中产生的非常大的那类空间来说,这就不是一个可接受的处理方法。下面所举的一个例子会表明,为什么假定可以把 Ω 的所有子集都考虑为事件是不合理的原因所在。从这个列表举出的另一个例子是,断言 $A \cap B = \varnothing$, 对应于断言,事件 A 和 B 是不相容的。在这个表中和后面,我们用简缩记法表示一个有穷集合数或无穷集合数的交集或并集:

$$\bigcap_{i=1}^{n} A_i = A_i \cap A_2 \cap \cdots \cap A_n$$

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n$$

$$\bigcap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \cdots$$

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \cdots$$

我们用同类符号表示和:

$$\sum_{i=1}^{n} P(A_i) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

此外,例如,如果 Ω 是一个有穷集,以及 f 是在 Ω 上定义的一个数值函数,那么,如果 $\Omega = \{a, b, c\}$,

$$\sum_{\omega \in \Omega} f(\omega) = f(a) + f(b) + f(c)$$

集合论符号和概率术语

133

集合论	**概率理论**
(a) $A \in \mathfrak{F}$	(a) A 是一个事件。
(b) $A \cap B = \varnothing$	(b) 事件 A 和 B 是不相容的。
(c) $A_i \cap A_j = \varnothing$,且 $i \neq j$	(c) 事件 A_1, A_2, $\cdots$, A_n 是两两不相容的。
(d) $\bigcap_{i=1}^{n} A_i = B$	(d) B 是事件 A_1, A_2, $\cdots$, A_n 一起发生时发生的事件。
(e) $\bigcup_{i=1} A_i = B$	(e) B 是在事件 A_1, A_2, $\cdots$, A_n 中至少有一个事件发生时发生的事件。
(f) $B = -A$	(f) B 是 A 不发生时发生的事件。
(g) $A = \varnothing$	(g) 事件 A 是不可能的。

(h) $A = \Omega$　　　　(h) 事件 A 是不可避免的或确定的。

(i) $B \subseteq A$　　　　(i) 如果事件 B 发生，那么 A 一定发生。

在(f)中，求补运算当然是相对于 Ω 的。比如说，如果 A 是在我们上面的例子中至少得到两次正面的事件，那么 $-A$ 是得到至少两次反面的事件，而且，

$$-A = \Omega - A = \{tht, tth, htt, ttt\}$$

事件的代数。为了使事件族有适当的集合论结构，我们必须为这个族假定某些闭合特性。通过假定，任何两个事件的并集也是一个事件，任何事件的补集也是一个事件，表示事件的集合族就成为通常集合论语言中的一个集合代数。通过另外假定，$\mathfrak{J}$ 的任何一个无穷序列元素的并集也在 $\mathfrak{J}$ 中，这个代数就成为一个 σ-代数。如果没有这些闭合性，那么在谈论自然应该有效的事件时会出现尴尬的局面。例如，如果事件 A 是第一次掷硬币时出现的正面，B 是第二次掷硬币时出现反面的事件，那么，不能考虑事件 $A \cap B$，是令人奇怪的，$A \cap B$ 是 A 和 B 联合发生的事件。σ-代数所要求的可数的并集，在直观上，不太明显，但是，这个特性通常是更深入的工作所需要的。在这里，为了讨论后面有这种特性或没有这种特性的那类事例，以及在何种程度上在更加复杂的语境中，它是一个在直观上可欲的特性，我们在这里会将它明确地包括进来。这些概念在下列定义中加以阐述。

定义 1. $\mathfrak{J}$ 是在 Ω 上的一个集合代数，当且仅当，$\mathfrak{J}$ 是 Ω 的一个非空子集族，而且，对于 $\mathfrak{J}$ 中的每个 A 和 B，

$$1.\ -A \in \mathfrak{J}$$

$$2.\ A \cup B \in \mathfrak{J}$$

此外，如果 $\mathfrak{J}$ 在可数的并集条件下是闭合的，即如果对于 A_1，

$A_2, \cdots, A_n, \cdots \in \mathfrak{I}$，$\bigcup_{i=1}^{\infty} A_i \in \mathfrak{I}$，那么 $\mathfrak{I}$ 是 Ω 上的一个 σ-代数。

关于集合代数的基本特性的下列定理的证明,作为练习留给读者。 134

定理 1. 如果 $\mathfrak{I}$ 是 Ω 上的一个集合代数,那么

(i) $\Omega \in \mathfrak{I}$,

(ii) $\varnothing \in \mathfrak{I}$,

(iii) 如果 $A \in \mathfrak{I}$ 和 $B \in \mathfrak{I}$，那么

$$A \cap B \in \mathfrak{I}$$

(iv) 如果 $A \in \mathfrak{I}$ 和 $B \in \mathfrak{I}$，那么

$$A - B \in \mathfrak{I}$$

(v) 如果 $\mathfrak{I}$ 也是一个 σ-代数,并且,$A_1, A_2, \cdots \in \mathfrak{I}$,那么,

$$\bigcap_{i=1}^{\infty} A_i \in \mathfrak{I}$$

定理 1 告诉了我们许多关于集合代数的结构。应该明显的是,一个集合代数可能在特征上是很平凡的。例如,如果 Ω 是我们计划在它上面定义概率的样本空间,那么我们就能选择只由 Ω 和空集组成的族 $\mathfrak{I}$ 作为一个集合代数。当然,显而易见的是,这种特殊的代数无论如何是无趣的。也很容易选择不是集合代数的某些直观上自然的集合族。在我们掷硬币的例子中,只由结果(有时被称为原子事件)的单位集合所组成的集合族,就不是一个集合代数,因为这些结果中的两个结果的并集本身不在集合族当中。[①] 再者,考虑这个例子,至少包含一个 h 的 Ω 的所有子集族也不是一个集合代数,因为最大的这种子集的补集本身不在这个族当中。

---

① 一个单位集合是只有单一元素作为它的惟一项的一个集合。

很容易从自然数构造出一个不是 σ-代数的代数例子。设 $\mathfrak{I}$ 是 Ω 的所有有穷子集加上这些有穷子集的补集。于是,偶数的集合不在这个代数中,尽管在 σ-代数中,因为它是在这个代数中的有穷集的可数的并集。

概率的公理。现在,我们准备定义概率空间。在这个定义中,我们假设了已经提到的 Ω、$\mathfrak{I}$ 和 P 的集合论结构:Ω 是一个非空集合,$\mathfrak{I}$ 是一个非空 Ω 的子集族,P 是在 $\mathfrak{I}$ 上的一个实值函数。这样,这个定义适用于有序的三元组(Ω, $\mathfrak{I}$, P),我们称之为集合函数结构,因为 P 是集合族上的一个函数。

定义 2. 一个集合函数结构 $\mathbf{\Omega} = (\Omega, \mathfrak{I}, P)$ 是一个有穷的相加概率空间,当且仅当,对于 $\mathfrak{I}$ 中的每个 A 和 B:

P1. $\mathfrak{I}$ 是在 Ω 上的一个集合代数;

P2. $P(A) \geqslant 0$;

P3. $P(\Omega) = 1$;

P4. 如果 $A \cap B = \varnothing$,那么 $P(A \cup B) = P(A) + P(B)$。

135 此外,如果也满足下列两个公理,$\mathbf{\Omega}$ 是一个概率空间(不限于有穷的相加性):

P5. $\mathfrak{I}$ 是 Ω 上的一个集合的 σ-代数;

P6. 如果 A_1, A_2, …是 $\mathfrak{I}$ 中的两两不相容事件的一个序列,即 $A_i \cap A_j = \varnothing$, $i \neq j$,那么

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

在 P5 和 P6 中表示的特性可能是独立的。例如,满足 P6 的概率测度 P 被说成是可数地相加的,而且,我们可能关心在不是 σ-代数的那些代数上的可数地相加的概率测度。另一方面,为了完全不受数学观点的影响,既满足 P5,也满足 P6,是合乎情

理的，而且，由于这种原因，我们已经假定，概率空间既满足 P5，也满足 P6。[①] 特别是，对于作为实验次数或样本大小趋于无穷大的渐近理论的标准发展来说，这种充分的自主性是根本的。如果集合 Ω 是有穷的，那么每个有穷相加概率空间也是一个概率空间，这应该是明显的。可能已经提到，概率空间有时需要一个附加的公设。这个公设断言，如果 $A \subseteq B$、$B \in \mathfrak{J}$ 和 $P(B)=0$，那么 $A \in \mathfrak{J}$。当一个概率空间满足这个附加公设时，它被称为是完备的。

现在，我们转向许多基本概念和定理。它们中的大多数对于有穷相加概率空间都成立，因此，根本不需要使用公理 P5 和 P6。

定理 2.

(i) $P(\varnothing)=0$；

(ii) 对于每个事件 A，$P(A)+P(-A)=1$；

(iii) 如果事件 A_i（其中，$1 \leqslant i \leqslant n$）是两两不相容的，那么

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)$$

证明：我们只证明(iii)，这要求用到数学归纳。对于 $n=1$，显然

$$P(A_1)=P(A_1)$$

我们对 n 的归纳假设是

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)$$

---

① 在本节的定义 14 之后给出的一个简单例子是，计算泊松分布的平均值。这种计算是针对概率的无穷序列的，而且需要公理 P5 和 P6。

现在考虑

$$P(\bigcup_{i=1}^{n+1} A_i)$$

136 既然

$$\bigcup_{i=1}^{n+1} A_i = (\bigcup_{i=1}^{n} A_i) \cup A_{n+1},$$

$$P(\bigcup_{i=1}^{n+1} A_i) = P((\bigcup_{i=1}^{n} A_i) \cup A_{n+1})$$

$$= P(\bigcup_{i=1}^{n} A_i) + P(A_{n+1}) \quad \text{通过 P4 和这个定理的假设}$$

$$= \sum_{i=1}^{n} P(A_i) + P(A_{n+1}) \quad \text{通过归纳假设}$$

$$= \sum_{i=1}^{n+1} P(A_i) \qquad \text{证毕}$$

离散的概率密度。我们接下来在一个集合 Ω 上定义离散的概率密度。在集合 Ω 上的这样一个离散的概率密度是类似于集合函数 P 的一个点函数。对于许多种应用来说,在所进行的详细计算中,运用一个离散的或连续的概率密度函数是基本的,或者至少是非常合乎情理的。

定义 3. 在一个非空集合 Ω 上的一个离散的概率密度是一个在 Ω 上定义的非负的实值函数 P,使得对于所有的 $\omega \in \Omega$ 来说,除了 Ω 的元素个数至多是可数的和 $\sum_{\omega \in \Omega} p(\omega) = 1$ 之外,$p(\omega) = 0$。

正如在下列定理中所表达的那样,总是能够用一个离散的概率密度,在已知的集合代数上,产生一个惟一的概率测度,省略其证明。

定理 3. 设 Ω 是一个非空集,设 p 是 Ω 上的一个离散的概

率密度，设 $\mathfrak{F}$ 是 Ω 上的一个 σ-代数，再设 P 是 $\mathfrak{F}$ 上定义的一个实值函数，使得如果 $A \in \mathfrak{F}$，那么 $P(A) = \sum_{\omega \in A} p(\omega)$。于是，$\boldsymbol{\Omega} = (\Omega, \mathfrak{F}, P)$ 是一个概率空间。

我们暂且推迟考虑无穷集上的连续密度。这里是四个更重要的离散密度。第一，存在着简单地掷一枚硬币的分布，通常称之为伯努利密度(Bernoulli density)，因为在这个概率空间 Ω 中，只有两个可能值，即只有两个元素。样本空间 $\Omega = \{0, 1\}$ 和 $p(1) = p$，$p(0) = q = 1 - p$，其中，$0 \leqslant p \leqslant 1$。[①]对于一枚公平硬币来说，$p = \frac{1}{2}$。第二，存在着一个密切相关的二项密度，它是从 n 次独立的伯努利试验中产生的。概率空间 $\Omega = \{0, 1, \cdots, n\}$，即概率空间是非负整数 $0 \leqslant k \leqslant n$ 的集合，并且，　137

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

这里，$\binom{n}{k}$ 是二项式系数：

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

二项密度的经典例子是，在每次试验中，把一枚硬币掷 n 次，出现一次正面的概率是 p。概率 $p(k)$ 是出现 k 次正面的概率。当 Ω 是一个数集时，通常谈到二项分布的频率和谈到二项密度

---

① 在这里和本书的许多其他地方使用的符号中，不断地容忍了语法的含混性。在语句“$\boldsymbol{\Omega} = (\Omega, \mathcal{F}, P)$ 是一个概率空间”中，表达式“$\boldsymbol{\Omega} = (\Omega, \mathcal{F}, P)$”是一个名词短语，但在语句“样本空间 $\Omega = \{0, 1\}$”中，主要动词是“=”。在每个特殊语境中，这种含混性是很容易解决的，如果不使用这种含混性，我们就需要使用太多冗长的明确阐述。

的频率是一样的，因此，我们有必要定义这个重要的联合的**分布**概念。设 f 是在一个非空的数值集合 Ω 上的一个离散密度。在 Ω 上定义

$$F(x) = \sum_{y \leqslant x} f(y)$$

于是，F 是根据 f 定义的(累积的)概率分布。注意当 Ω 是一个任意的有穷集合时，分布概念不是很明确定义的。为了超越数值集合进行归纳，有必要对 Ω 的元素进行一种有意义的线性排序。

第三，存在着含有参数 λ 的**泊松**密度，这里 $\lambda \geqslant 0$。①概率空间 $\Omega = \{0, 1, 2, \cdots\}$ 和

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

泊松密度或分布的应用补充了二项分布的那些应用。在应用二项分布时，我们数出某些现象发生的次数，比如，在给定的试验中，出现正面的次数，或者，在给定的数据中，出现一种疾病的次数。这样，例如，第一次试验中可以出现一次正面，但在第二次试验中不出现。相反，还有我们感兴趣的许多别的现象是，在一个给定的时间间隔内(例如，一个小时、一天、一年)，共发生多少次。其他例子是，在二十四小时内通过一个特定的交换机呼叫过的电话号码，一周之内一个机场降落的航班数，以及两分钟之内盖革计数器的放电总数。泊松分布也能被应用于空间现象；比如，从一个城市的饮用水中提取一升水样，在这个水样中发现的特定种类的细菌数。在这些空间类型或时间类型的例子中所

① 泊松(Poisson，1781—1840)是19世纪伟大的法国数学物理学家之一。他推导的泊松密度发表于1837年，但直到19世纪末，才真正被公认为是一项有意义的贡献。

隐含的是进行连续地采样。比如,我们可以查看连续二百天的电话交换数据。尽管在现实的时间间隔内或空间范围内所描述的现象中,还没有一种现象发生的有限次数会很大,比如说,10^{100},但在数学上定义每个非负整数的密度,是很方便的。当然,一种理由是,根本没有自然的停止点,但这肯定不是主要理由。选择一个更能适当地研究实际观察的有限范围总是很容易的,但当所用的这个有限范围很大时,确定这种分布的数学特性或理论特性就会更困难。

第四,存在着含有参数 p 的几何密度,其中,$0 \leqslant p \leqslant 1$。概率空间 $\Omega = \{1, 2, 3, \cdots\}$ 和 138

$$p(k) = p(1-p)^{k-1}$$

一个理论上的重要例子是,在具有成功概率 p 的连续的伯努利试验中,第一次得到正面的试验次数 k 的密度。例如,在投掷一枚重心有偏向的硬币时,得到正面的概率为 $p=0.25$,投掷到第五次时第一次出现正面的概率是 0.079(近似值)。正如我们在关于概率的倾向性解释那一节将会看到的那样,当使用离散试验时,几何密度是指在研究放射性衰变时用到的自然密度。

条件概率。我们接下来考虑条件概率的重要概念。它的核心作用来自这样的事实:用条件术语表达依赖性尤其是时间依赖性。

定义 4. 如果 A 和 B 在 $\Im$ 中,并且,$P(B) > 0$,那么

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

我们把 $P(A|B)$称为在已知事件 B 发生的条件下事件 A 的**条件概率**。作为一个例子,考虑把一枚硬币掷两次的情境。$\Omega = \{hh, ht, th, tt\}$,$\Im$ 是 Ω 的所有子集的集合,而且,对于在 Ω 中

的所有 ω，$p(\omega)=\frac{1}{4}$。有两个事件 A 和 B，使得 A 是在两次投掷中只出现一次正面的事件，B 是两次投掷中至少出现一次正面的事件。显然，$P(A)=\frac{1}{2}$，而 $P(B)=\frac{3}{4}$。因为 $A\cap B=A$，$P(A\cap B)=P(A)=\frac{1}{2}$。因此，$P(A\mid B)=\frac{P(A)}{P(B)}=\frac{\left(\frac{1}{2}\right)}{\left(\frac{3}{4}\right)}=\frac{2}{3}$。

由于显而易见的理由，关于条件概率的下列定理通常被贴上了乘法定理的标签。

定理 4. (乘法定理)。如果 A_1，A_2，A_3，…，A_n 在 $\mathfrak{J}$ 中，而且，$P(\bigcap_{i=1}^{n}A_i)>0$，那么

$$P(\bigcap_{i=1}^{n}A_i)=P(A_1)\cdot P(A_2\mid A_1)\cdot P(A_3\mid\bigcap_{i=1}^{2}A_i)\cdots P(A_n\mid\bigcap_{i=1}^{n-1}A_i)$$

证明： 我们运用对于 n 的归纳。对于 $n=1$，显然，$P(A_1)=P(A_1)$。我们对 n 的归纳假设是

$$P(\bigcap_{i=1}^{n}A_i)=P(A_1)\cdots P(A_n\mid\bigcap_{i=1}^{n-1}A_i) \tag{1}$$

但是，根据条件概率的定义

$$P(\bigcap_{i=1}^{n+1}A_i)=P(\bigcap_{i=1}^{n}A_i)\cdot P(A_{n+1}\mid\bigcap_{i=1}^{n}A_i)$$

因此，运用(1)

$$P(\bigcap_{i=1}^{n+1}A_i)=P(A_1)\cdot P(A_2\mid A_1)\cdots P(A_n\mid\bigcap_{i=1}^{n-1}A_i)\cdot P(A_{n+1}\mid\bigcap_{i=1}^{n}A_i)$$

证毕。

乘法定理的应用通常包括其他简化的假设。马尔可夫链提供了好的事例。从直观上看，在一系列试验中，一次试验所发生的概率依赖于前一次试验所发生的情况，与较早的试验无关，这就是一个(一阶)马尔可夫链。(在这一点上没有给出形式定义。①)一个稍微人为的例子会提供一个简单的说明。设想，假如我们说，记住一个特定的算法，只依赖于在最后一次学习试验中是否能够记住它，与较早的学习试验无关。设 r 表示记住，f 表示忘记。于是，转移矩阵是： 139

$$\begin{array}{c} \\ r \\ f \end{array}\begin{array}{c} \begin{array}{cc} r & f \end{array} \\ \begin{bmatrix} 0.75 & 0.25 \\ 0.10 & 0.90 \end{bmatrix} \end{array}$$

这里，两行表示第 $n-1$ 次试验的态，两列表示第 n 次试验的可能态。因此，如果在第 $n-1$ 次试验中，记住算法，那么在第 n 次试验中记住这个算法的概率是 0.75。假定我们拥有的观察序列是 $ffffrrfrrr$，再进一步设想，在第一次试验中，忘记的概率是 1.0——根据第一次出现。然后，我们能用这些特殊假设和乘法定理计算我们所观察的序列的概率。

$$\begin{aligned} & p(ffffrrfrrr) \\ & = p(f)p(f \mid f)^3 p(r \mid f)p(r \mid r)p(f \mid r)p(r \mid f)p(r \mid r)^2 \\ & = (1)(0.9)^3(0.10)(0.75)(0.25)(0.10)(0.75)^2 \\ & = 0.000\,768\,9 \qquad \text{(近似值)} \end{aligned}$$

下面的两个定理表明，条件概率表现得像普通概率一样。

---

① 马尔可夫链的一种详细的形式处理，在第 7 章的最后一节给出，这一节论述了关于因果性过程的可逆性。在 4.5 节中，提供了一个不太详尽的进展，但强调了各态历经特性。

定理 5. 如果 A、B 和 C 在 $\mathfrak{F}$ 中，并且，如果 $P(A)>0$，那么：

(i) $P(B \mid A) \geqslant 0$；

(ii) $P(\Omega \mid A)=1$；

(iii) $P(A \mid A)=1$；

(iv) 如果，$B \cap C=\varnothing$，那么 $P(B \cup C \mid A)=P(B \mid A)+P(C \mid A)$；

(v) 如果 $P(B)>0$，那么 $P(A \mid B)=P(A) \cdot \dfrac{P(B \mid A)}{P(B)}$；

(vi) 如果 $A \cap B=\varnothing$，那么 $P(B \mid A)=0$，

(vii) 如果 $A \subseteq B$，那么 $P(B \mid A)=1$。

定理 6. 如果 $A \in \mathfrak{F}$ 和 $P(A)>0$，那么$(\Omega, \mathfrak{F}, P_A)$是一个有穷相加概率空间，使得对于在 $\mathfrak{F}$ 中的每个 B，$P_A(B)=P(B \mid A)$。此外，如果$(\Omega, \mathfrak{F}, P)$是可数地相加的，那么$(\Omega, \mathfrak{F}, P_A)$也是如此。

证明：已知 $\mathfrak{F}$ 是 Ω 上的一个代数。根据定理 5(i)，我们有 $P_A(B) \geqslant 0$，它满足定义 2 的 P2。根据定理 5(ii)，我们有 $P_A(\Omega)=1$，它满足 P3。根据定理 5(iv)，如果 $B \cap C=\varnothing$，那么 $P_A(B \cup C)=P_A(B)+P_A(C)$，它满足 P4，而且，同样符合可数的相加性。

证毕。

140 定理 6 令人感兴趣的地方是，它表明，我们如何可以用一个事件的知识产生一个新的概率空间。

定理 7.（关于总概率的定理）。如果 $\Omega=\bigcup_{i=1}^{n} A_i$，而且，如果 $A_i \in \mathfrak{F}$，$P(A_i)>0$ 和 $A_i \cap A_j=\varnothing$，其中，$i \neq j$，$1 \leqslant i, j \leqslant n$，那么对于每个 $B \in \mathfrak{F}$，

$$P(B)=\sum_{i=1}^{n}P(B\mid A_i)P(A_i)$$

证明：运用这个定理的假设和逻辑分布定律，我们有

$$B=B\cap\Omega=B\cap(\bigcup_{i=1}^{n}A_i)=\bigcup_{i=1}^{n}(B\cap A_i)\qquad(2)$$

因此，

$$P(B)=P(\bigcup_{i=1}^{n}(B\cap A_i))\qquad(3)$$

因为通过假设 $A_i\cap A_j=\varnothing$，其中，$i\neq j$，我们可以把关于有穷相加定理（定理 2(iii)）应用到(3)，然后得到

$$P(B)=\sum_{i=1}^{n}P(B\cap A_i)\qquad(4)$$

但通过定义 4

$$P(B\cap A_i)=P(B\mid A_i)P(A_i)\qquad(5)$$

然后从(4)和(5)可立即推出我们的定理。

证毕

我们以一种基本方式运用定理 7 证明下一个定理，尽管它的证明很简单，但它在概率理论中占有重要的地位。在陈述这个定理时，把 H_i 看成是假设，但这显然不是形式上的要求。

定理 8.（贝叶斯定理）。[1] *如果* $\bigcup_{i=1}^{n}H_i=\Omega$，$P(H_i)>0$，$H_i\cap H_j=\varnothing$，*其中*，$i\neq j$ *和* $1\leqslant i,j\leqslant n$，*如果* $B\in\mathfrak{J}$ *和* $P(B)>0$，*那么*

$$P(H_i\mid B)=\frac{P(B\mid H_i)P(H_i)}{\sum_{j=1}^{n}P(B\mid H_j)P(H_j)}$$

---

[1] 受人尊敬的托马斯·贝叶斯(Thomas Bayes 1763)撰写了关于这个定理及其使用的一本重要的早期论文集。

证明： 根据定理 5(v)

$$P(H_i \mid B) = \frac{P(B \mid H_i)P(H_i)}{P(B)} \tag{6}$$

但是，通过定理 7，

$$P(B) = \sum_{j=1}^{n} P(B \mid H_j)P(H_j) \tag{7}$$

把(7)代入(6)，我们立即获得所需要的结果。 证毕

正如我们在主观概率那一节将会看到的那样，这个数学上平凡的定理有着广泛的应用。对贝叶斯定理的术语值得作出某种评论。H_i 通常是一个互相排斥的和穷举的假设集合，这个集合中的一个项提供了对事件 B 的相对最好的说明。概率 $P(H_i)$被称为假设 H_i 的**先验**概率。这些先验概率是关于应用概率理论的许多争论的核心。条件概率$P(B|H_i)$被称为 H_i 在 B 上的**可能性**，或者 H_i 在证据 B 上的可能性。条件概率 $P(H_i|B)$被称为，在已知所观察到的事件 B 的条件下，假设 H_i 的**后验**概率。贝叶斯定理的数学地位是无可争议的，这应该是可理解的。在无可争辩的意义上，这个定理是定义 2 的公理的一个逻辑推论。

141

贝叶斯公设。这个公设是把所有的先验可能性 $P(H_i)$都看成是相等的经典程序。重要的是注意到，这个公设只关注概率理论的*应用*；与概率空间的纯数学理论无关。需要强调的是，贝叶斯定理在逻辑上独立于贝叶斯公设。同样独立于贝叶斯公设的，正如上面所提供的那样，是**贝叶斯的行为规则**：接受最大后验概率$P(H_i|B)$的假设。[1] 当然，在应用这个规则时，我们有

[1] 然而，这相当于说，实践比所推荐的行动更多地报告了后验概率，推荐的行动可能依赖于其他因素，比如，效用或实践。

时用贝叶斯公设计算先验概率 $P(H_i)$，它是计算后验概率 $P(H_i|B)$所需要的。

应用贝叶斯公设的一个例子。 假定有一个装有四个球的罐子，我们只知道下列之一是真的：

H_1：四个球都是白的；

H_2：两个球是白色的，两个球是黑色的。

进一步假定从罐子里拿出一个白球。令人感兴趣的问题是，确立一个适合于计算 H_1 和 H_2 的后验概率的概率空间。

在概率理论中，根本没有通用的规则来构造可能结果的集合 Ω，尽管概率理论确实表明，Ω 一旦得到了详细的说明，就会如何操纵它。产生概率空间的某些标准会在后面考虑。当前，无论如何，能够以这样一种方式构造 Ω：连续地从一个罐子里抽出一个球，无可替代地连续抽取四次，将会必定惟一地确定 H_1 或 H_2 哪个假设是真的。这样，

$$\Omega=\{wwww, wwbb, bbww, wbbw, bwwb, wbwb, bwbw\}$$

我们会观察到，H_1 和 H_2 是 Ω 的子集。

$$H_1 = \{wwww\}$$

$$H_2 = \{wwbb, bbww, wbbw, bwwb, wbwb, bwbw\}$$

$$=- H_1$$

在抽样时，第一次抽出一个白球的事件，用 W 表示。

$$W = \{wwww, wbwb, wwbb, wbbw\}$$

贝叶斯公设告诉我们，这两个假设的概率是先天地相等的。也 142
就是

$$P(H_1) = \frac{1}{2}$$

$$P(H_2)=\frac{1}{2}$$

通过考虑每个假设中白球占总球数之比，能够计算出 H_i 关于事件 W 的可能性：

$$P(W\mid H_1)=1$$

$$P(W\mid H_2)=\frac{1}{2}$$

可能性的这种计算是以我们所称的简单随机抽样公设为基础的。这个公设的直观思想是，如果 H 是真的假设，并且，如果 $\omega_1,\omega_2\in H$（即，ω_1 和 ω_2 是样本空间 Ω 的元素，也是子集 H 的元素），那么

$$P(\{\omega_1\})=P(\{\omega_2\})$$

（既然$\{\omega_1\}$或$\{\omega_2\}$不可能在 $\mathfrak{I}$ 中，所以，这不是对该公设的精确陈述。）换句话说，我们用一个抽样程序，使得如果 H 是真的假设，那么与 H 相容的任何一种可能的抽样结果，都与任何其他的抽样结果一样可能。重要的是注意到，在许多情况下，我们的实验程序允许我们适当灵活地保证满足这个公设。另一方面，这样的控制通常根本不能应用于先验概率，因为先验概率是我们在抽样之前的信息的函数。因此而产生了关于贝叶斯公设的争论。

每个假设的后验概率都能通过定理 8 来计算，就这个例子而言，后验概率的计算能被陈述如下：

$$P(H_1\mid W)=\frac{1\cdot\frac{1}{2}}{1\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}}=\frac{2}{3}$$

$$P(H_2 \mid W) = \frac{\frac{1}{2} \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}} = \frac{1}{3}$$

这是以一种明显的方式代入合适的公式得出的。

我们可以构造另一个空间,这个空间产生出的结果与前面例子中的 Ω 一样。在 Ω 的定义中产生的困难不是概率理论固有的,而是表明了在获得对概率空间的一种实际描述时的模糊性。在上述例子中,显然,我们只基于最多三次独立的抽样,就能够决定哪一个直观假设是正确的。因此,我们可以使用

$$\Omega' = \{www, wwb, wbw, bww, bbw, bwb, wbb\}$$

可以看到,在 Ω' 中的元素个数等于 Ω 中的元素个数。也就是说,它们的基数是相同的。同样,H_1 和 H_2 被指定为

$$H_1 = \{www\}$$

$$H_2 = \{wwb, wbw, bww, bbw, bwb, wbb\} = -H_1$$

只要略加考虑就足以看出,Ω 和 Ω' 在所有其余方面都是相同的,而且,H_1 再一次具有较高的后验概率。 143

如果用有序的 n 元组的替代集合来表示 Ω 或 Ω' 的元素,显然会出现困难。我们于是有

$$\Omega'' = \{\{w\}, \{w, b\}\}$$

而且,H_1 和 H_2 将成为

$$H_1 = \{\{w\}\}$$

$$H_2 = \{\{w, b\}\}$$

适合这个例子的代数是

$$\mathfrak{F}' = \{\varnothing, \{\{w\}\}, \{\{w, b\}\}, \Omega''\}$$

于是问题是,$\mathfrak{J}$ 的什么子集是 W,即得到一个白球的事件？显然,当 Ω''是基本空间时,不可能适当地描述 W 的特征。

然而,Ω 和 Ω' 没有代表惟一合法的可能性。由于几种原因,许多统计学家喜欢下列空间：

$$\Omega''' = \{(4, w), (2, w), (2, b)\}$$

因此,

$$H_1 = \{(4, w)\}$$

$$H_3 = \{(2, w), (2, b)\}$$

每个有序对的第一项的数字都表明了在特定假设(H_1 或 H_2)的条件下罐子里的白球总数。第二项表明了抽出一个球的一种可能结果。注意在 Ω''' 中只有三个元素,这是优越于 Ω 和 Ω' 的地方。喜欢 Ω''' 的第二个原因是,它表明了只抽样一次的可能结果,这里,Ω 的元素表明了无可替代地抽样四次的结果,Ω' 的元素表明了无可替代地抽样三次的结果。因此,Ω 和 Ω' 中的元素总数过多。空间 Ω''' 切合这个问题,Ω 和 Ω' 则并非如此。

无论用什么样的样本空间,关键之处在于接受贝叶斯公设。许多专家,尤其是拥护概率的客观解释的那些专家,对这个公设的合理性持有批评态度。不幸的是,没有贝叶斯公设或产生先验概率的某个类似的公设,就不可能直接应用前面陈述的贝叶斯的行为规则。

应对这个困难的一个建议是下列原理,即**最大可能性原理：接受所观察事件的可能性是最大的假设。**H_j 是可接受的,当且仅当,

$$P(E \mid H_j) = \max_i P(E \mid H_i)$$

作为一个明显的结果，我们有：[1]

定理 9. 如果假设 H_i 的总数是有穷的，那么贝叶斯的行为规则与贝叶斯公设一起导致了把相同的假设接受为最大可能性原理。 144

一个实验的概念。我们接下来转向一个实验的概念。我们相对于样本空间 Ω，把一个实验定义为 Ω 的任何一种划分(partition)，即 Ω 的非空子集的任何一种集合，使得：(i) 这种集合是集合代数 $\mathfrak{J}$ 的一个子集，(ii) 在这种集合中，任何两个不同集合的交集都是空的，以及(iii) 这种集合的并集是 Ω 本身。就我所知，对一个实验的这种定义是由柯尔莫哥洛夫首创的(1933/1950)。需要注意的是，这个定义是概率理论中的一个技术概念，并不意味着捕获到了在这个术语的一般科学用法中所蕴含的含义的许多细微差别。尽管如此，很容易举例说明，这里所定义的技术概念对应于我们关于实验数据的许多直观思想。考虑从记录多种数据的人造卫星获得的复杂的连续的时间记录。完整的样本空间 Ω 把所研究的时期和变量的所有可能的观察记录作为它的元素。只对太阳黑子的数据感兴趣的一位气象学家，根据且只根据太阳黑子的数据，通过划分可能观察到的记录来定义关于 Ω 的一个实验。在这种数据简化中，忽略了关于其他任何类型的变量数据。在把一个定量理论任意详尽地应用到一个实验结果时，这样的简化通常是必要的。

独立性。现在，我们定义实验的彼此独立的重要概念。

定义 5. n 次实验 $M^{(1)}, \cdots, M^{(n)}$ 是(彼此)独立的，当且仅当，对于每个 $A_i \in M^{(i)}$（其中，$1 \leqslant i \leqslant n$），

---

[1] 至于把这个定理推广到无穷集，可参见在无穷小框架内阐述的，即非标准分析的第 5 节的定理 3。

$$P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$$

注意，A_1 是在 $M^{(1)}$ 中的任何一个集合，以此类推，$\prod_{i=1}^{n}$ 是 n 个概率的算术积。

n 个事件 $A_1, A_2, \cdots, A_n$ 是(**彼此**)**独立的**，当且仅当，n 次实验 $M^{(1)}, M^{(2)}, \cdots, M^{(n)}$ 是彼此独立的，在这种情况下，

$$M^{(i)} = \{A_i, -A_i\}$$

我们考虑举两个例子来说明独立性概念。

第一个例子。设想把一枚硬币连续掷三次，对于这种情况来说，

$$\Omega = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$$

设 $\mathfrak{I}$ 是 Ω 的所有子集的集合，而且，对于所有的 $\omega \in \Omega$ 来说，设 $p(\omega) = \frac{1}{8}$。后面的这个条件确保这一枚硬币是公平的。把实验 $M^{(1)}$ 定义如下：对于 $i = 0, 1, 2, 3$, $A_i \in M^{(1)}$，当且仅当，A_i 是恰好得到 i 个正面的事件。我们能够轻易地证实，$M^{(1)}$ 是一个实验，因为 $M^{(1)}$ 显然是 Ω 的一种划分。同样，对于 $i = 0, 1, 2$，定义实验 $M^{(2)}$，使得 $B_i \in M^{(2)}$，当且仅当，B_i 是正好 i 次交替得到正面和反面的事件。

145

问题是要确定，$M^{(1)}$ 和 $M^{(2)}$ 是否是彼此独立的实验。有 12 种情况可供考虑，因为 $M^{(1)}$ 有四个元素，$M^{(2)}$ 有三个元素。然而，以 $i = 0$ 的情形为例。在这种情形中，$A_0 = \{ttt\}$，为此 $P(A_0) = \frac{1}{8}$，以及 $B_0 = \{hhh, ttt\}$，为此 $P(B_0) = \frac{1}{4}$。显然，

$$P(A_0 \cap B_0) = P(A_0) = \frac{1}{8}$$

但是，

$$P(A_0)\cdot P(B_0)=\frac{1}{32}$$

因此，实验 $M^{(1)}$ 和 $M^{(2)}$ 不是独立的。

第二个例子。设想同时把两枚硬币掷两次。一枚硬币是红的，一枚硬币是黑的。红硬币的正面用"H"表示，黑硬币的正面用"h"表示，等等。就这个例子而言，概率空间是

$$\Omega=\{HhHh,\ HhHt,\ HhTh,\ \cdots,\ TtTt\}$$

像以前一样，设 $\mathfrak{I}$ 是 Ω 的所有子集的集合，并且，对于所有的 $\omega\in\Omega$, $p(\omega)=\frac{1}{16}$。对于 $i=0,1,2$，描述一个实验 $M^{(1)}$，其中，$A_i\in M^{(1)}$，当且仅当，A_i 正好是 i 次得到红硬币正面的事件。对于 $i=0,1$，定义实验 $M^{(2)}$，使得 $B_i\in M^{(2)}$，当且仅当，B_i 正好是 i 次交替得到黑硬币的正面和反面的事件。当然有六种不同的情况可供考虑。以 A_0 和 B_1 为例。于是

$$A_0=\{ThTh,\ ThTt,\ TtTh,\ TtTt\}$$

为此 $P(A_0)=\frac{1}{4}$，以及

$$B_1=\{HhHt,\ HtHh,\ HtTh,\ HhTt,\ ThHt,\ TtHh,\ TtTh,\ ThTt\}$$

为此 $P(B_1)=\frac{1}{2}$，此外，

$$A_0\cap B_1=\{TtTh,\ ThTt\}$$

$$P(A_0\cap B_1)=\frac{1}{8}$$

$$P(A_0)\cdot P(B_1)=\frac{1}{4}\cdot\frac{1}{2}=\frac{1}{8}$$

对其他五种情况进行相似的计算确定了,实验 $M^{(1)}$ 和 $M^{(2)}$ 是彼此独立的。

下面的定理是显而易见的,而它的逆定理是错误的。

定理 10. ***如果 n 次实验是彼此独立的,那么,对于 $m<n$,它们的任何 m 个实验都是彼此独立的。***

显然,事件 $A_1, A_2, \cdots, A_n$ 中的两两事件的独立性,不能确保这些事件是彼此独立的。考虑下列例子[归功于伯恩斯坦(S. Bernstein)]。设 $\Omega=\{a_1, a_2, a_3, a_4\}$ 和 $A_1=\{a_1, a_2\}$、$A_2=\{a_1, a_3\}$ 以及 $A_3=\{a_1, a_4\}$。对于 Ω 中的每个 ω,$p(\omega)=\frac{1}{4}$。容易计算,

$$P(A_1)=P(A_2)=P(A_3)=\frac{1}{2}$$

146

$$P(A_1\cap A_2)=P(A_2\cap A_3)=P(A_1\cap A_3)=\frac{1}{4}$$

这确立了两两事件的独立性。但相反,

$$P(A_1\cap A_2\cap A_3)=\frac{1}{4}\text{, 而不是}\frac{1}{8}\text{,}$$

正如所预期的那样。这三个事件是两两独立的,但是,它们不是彼此独立的。①

在两个事件 A 和 B 的重要情况下,从前面的定义容易表明,它们是独立的,当且仅当,

---

① 关于另外一个具有物理学根据的例子,直接源于量子纠缠理论,参见第234页脚注①。

$$P(A \cap B) = P(A)P(B)$$

独立性概念是概率理论和一般科学的最深刻与最基本的思想之一。精确定义的独立性的概率概念很好地表达了支持一般概念的直观的核心思想。这两种最基本的思想所关注的是,这里和现在所发生的事件独立于遥远的过去或很远的地方发生的事件。关于存在着空间或时间距离上的直接因果关系的任何一种主张都会受到怀疑,并且,通过中间的近因,努力为这样一种超距作用作出辩解。[①] 独立性的概率概念以相当一般的方式为分析这些关于因果关系的思想提供了一种核心手段(Suppes, 1970a)。

随机变量。现在,我们转向重要的随机变量概念。在科学中的概率理论的许多应用中,潜在的样本空间具有很复杂的集合论结构,可是,在任意特定的时刻,我们只对这个空间的某些方面感兴趣。为了达到这个目的,我们在样本空间上定义某些实值函数,并研究这些函数的概率行为。

为了与统计学中的常用符号相一致,我们因而用黑体字“$\mathbf{X}$”、“$\mathbf{Y}$”等表示随机变量。

定义 6. $\mathbf{X}$ 是一个随机变量,当且仅当,$\mathbf{X}$ 是一个实值函数,它的定义域是概率空间 Ω,而且,对于每个实数 x,集合 $\{\omega: \mathbf{X}(\omega) < x\}$ 在 $\mathfrak{I}$ 中。

$\{\omega: \mathbf{X}(\omega) < x\}$ 在 $\mathfrak{I}$ 中的必要条件,根本不限于有穷概率空间,对于有穷概率空间来说,$\mathfrak{I}$ 是所有子集的集合。在下面的定义中,一个随机变量如何可以表示一次实验,变得更加明确。

定义 7. 设 M 是一次实验,再设 $\mathbf{X}$ 是在 Ω 上定义的一个实

① 并不总是这样。在 18 世纪,引力是超距作用的主要事例,在 19 世纪也有令人尊敬的电和磁的超距理论。

值函数。于是,$\boldsymbol{X}$ 是关于实验 M 的一个随机变量,当且仅当,对于 M 中的每个 A 来说,有一个实数 k_A,使得对于 A 中的每个 ω, $\boldsymbol{X}(\omega) = k_A$。

在讨论一个非数值的离散样本空间时,比如,

$$\Omega = \{hh, ht, th, tt\}$$

147 很自然建议用下列数值样本空间取而代之

$$\Omega_1 = \{0, 1, 2\}$$

在这里,每个整数表示把一枚硬币掷两次时可能出现正面的个数。如所期望的那样,一种可替代的建议是,在 Ω 上定义一个随机变量,而不是用 Ω_1 替代 Ω。因此,我们可以在 Ω 上定义这个随机变量,使得 $\boldsymbol{X}(hh) = 2$、$\boldsymbol{X}(ht) = 1$、$\boldsymbol{X}(th) = 1$ 和 $\boldsymbol{X}(tt) = 0$,而且,随机变量 $\boldsymbol{X}$ 现在达到了最初为 Ω_1 所建议的目标。正如已经评论的那样,用黑体大写字母"**X**"、"**Y**"、"**Z**"表示随机变量是统计学中的习惯。另一方面,不习惯于写下是一个随机变量的自变量,即在统计学文献中,遇到像"$\boldsymbol{X}(hh)$"之类的表示,是令人惊奇的和异常的。为了与统计学的常用简写符号相一致,我们将引入某个缩写符号。

定义 8. 设 $\boldsymbol{X}$ 在 Ω 上的一个随机变量,于是,

$$P(\boldsymbol{X} \leqslant x) = P(\{\omega: \boldsymbol{X}(\omega) \leqslant x\})$$

和

$$P(a \leqslant \boldsymbol{X} \leqslant b) = P(\{\omega: a \leqslant \boldsymbol{X}(\omega) \leqslant b\})$$

可以为符号给出相对应的定义: $P(\boldsymbol{X} = x)$、$P(\boldsymbol{X} < x)$、$P(\boldsymbol{X} \geqslant a)$ 和 $P(\boldsymbol{X} > x)$。(注意,我们用字母"a"、"b"、"x"作为数值变量。)

对于把一枚硬币掷两次的情况来说,假定这枚硬币是公平

的,即对于 $\omega \in \Omega$, $p(\omega)=\frac{1}{4}$, 我们有

$$\begin{aligned}
P(\boldsymbol{X} \leqslant 1) &= P(\{\omega: \boldsymbol{X}(\omega) \leqslant 1\}) \\
&= P(\{HT, TH, TT\}) \\
&= P\{HT\} + P\{TH\} + P\{TT\} \\
&= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \\
&= \frac{3}{4}。
\end{aligned}$$

同样,我们容易推出

$$P(\boldsymbol{X} > 0) = \frac{3}{4}$$

$$P(\boldsymbol{X} < 1) = \frac{1}{4}$$

很容易用这个符号把一个随机变量 $\boldsymbol{X}$ 的**分布** F 定义为:

$$F(x) = P(\boldsymbol{X} \leqslant x)$$

一个随机变量的分布 F 有时也被称为**累积**分布。注意,随机变量总是有分布的,与一个任意的有穷集合 Ω 相反,有穷集合 Ω 可能有在它之上定义的一个密度,而不是一种分布,因为 Ω 的元素没有相关的或自然的排序,这一点先前已明确。既然一个随机变量的值域是一个数的集合,所以,不会产生定义这种分布的排序问题。

尽管这种分类不是穷举的,但是,对离散随机变量和连续随机变量作出分类是有益的。这个定义不允许我们从一个随机变量是离散的事实推论出潜在的概率空间 Ω 也是离散的结论。 148

定义 9. 一个随机变量 $\boldsymbol{X}$ 是离散的,当且仅当,$\boldsymbol{X}$ 的值的集

合是可数的。

换言之,一个离散变量是一个函数,它的值域是一个可数集。一个离散变量的离散密度是很容易定义的:

定义 10. 设 $\mathbf{X}$ 是一个离散的随机变量。于是,$\mathbf{X}$ 的离散密度是函数 p,使得对于 $\mathbf{X}$ 的任何一个值 x,

$$p(x) = P(\mathbf{X} = x)$$

在把一枚公平的硬币掷两次的实验中,数值随机变量 $\mathbf{X}$ 的值表示两次投掷中获得正面的个数,$\mathbf{X}$ 的离散密度 p 是:

$$p(0) = \frac{1}{4}$$

$$p(1) = \frac{1}{2}$$

$$p(2) = \frac{1}{4}$$

注意,如何解释"一个随机变量 $\mathbf{X}$ 的密度"这个短语。这个密度实际上在这个随机变量的可能值的集合上,而且,密度 f 或分布 F 可能总是根据随机变量 $\mathbf{X}$ 和在基本的样本空间上的概率测度 P 下定义的。

分段连续的密度和分布。在下面的定义中,我们用了分段连续函数的概念。这是一个连续函数,除了在有穷数目的点之外,而且,在每个不连续的点,都存在着右极限和左极限。

定义 11. 一个随机变量 $\mathbf{X}$ 是分段连续的,当且仅当,有一个分段连续密度 f,使得对于每两个实数 a 和 b(其中,$a \leqslant b$),

$$P(a \leqslant \mathbf{X} \leqslant b) = \int_a^b f(x)dx$$

通常情况下,我们总是说,一个分段连续随机变量的密度是什

么。定义 11 的概括在数学文献中是重要的,但这里将不考虑这个问题。

再一次注意一个随机变量 $\boldsymbol{X}$ 的密度 f 和它的分布 F 之间的明显关系。

$$F(x)=\sum_{y\leqslant x}f(y) \qquad (离散型)$$

$$F(x)=\int_{-\infty}^{x}f(x)dx \qquad (分段连续型)$$

从这些方程来看,为什么一种分布 F 也被称为一种累积分布, 149 是显而易见的。

下面是在非数值样本空间上的一个连续随机变量的例子。(当这个随机变量有一个连续的不只是分段连续的密度时,我们将把它简单地说成是连续的。)设想为检验新型火箭的精确性所设计的一个实验。火箭瞄准三千里之外的一个目标,实验是要确定这个火箭实际上落在哪里。我们的基本的样本空间是地面上点的集合。这里有趣的一个数值随机变量是,对于实验的任何一种结果来说,这个变量的值是火箭的目标和实际位置之间的距离。

随机变量的期望值。现在,我们转向定义随机变量的几个基本特性。这些定义是为离散随机变量和分段连续随机变量提供的。我们从一个随机变量的期望值或平均值开始。

定义 12. 设 $\boldsymbol{X}$ 是一个具有密度 f 的离散的或分段连续的随机变量。于是,倘若存在着和或积分的话,通过下列等式给出 $\boldsymbol{X}$ 的期望值或平均值 $E(\boldsymbol{X})$:①

$$E(\boldsymbol{X})=\sum xf(x) \qquad (离散型)$$

---

① 这个和被认为是这个随机变量的可能数值的集合。

$$E(\boldsymbol{X}) = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{（分段连续型）}$$

应该清楚的是，一个随机变量的期望值不一定存在。例如，柯西密度的期望值

$$f(x) = \frac{1}{\pi(1+x^2)}$$

是不存在的。

还需要把期望值的定义推广到一个随机变量的任何一个连续的数值函数。

定义 13. 设 $\boldsymbol{X}$ 是一个具有密度 f 的离散的或分段连续的随机变量，再设 h 是在后一种情况下的一个连续函数。于是，倘若存在着和或积分的话，通过下列等式给出 $h(\boldsymbol{X})$ 的期望值或平均值 $E(h(\boldsymbol{X}))$：

$$E(h(\boldsymbol{X})) = \sum h(x) f(x) \qquad \text{（离散型）}$$

$$E(h(\boldsymbol{X})) = \int_{-\infty}^{\infty} h(x) f(x) dx \qquad \text{（分段连续型）}$$

我们可以用定义 14 来定义关于一个随机变量的原点矩和平均值矩。

定义 14. 设 $\boldsymbol{X}$ 是一个具有密度 f 的离散的或分段连续的随机变量。于是，倘若期望值存在的话，那么：

150 (i) $\mu = \mu_{\boldsymbol{X}} = \mu_1' = E(\boldsymbol{X}) = \boldsymbol{X}$ 的平均值，

(ii) $\mu_r' = \mu_r(\boldsymbol{X}) =$ 关于 $\boldsymbol{X} = E(\boldsymbol{X}^r)$ 的原点的 r 阶矩，

(iii) $m_r = m_r(\boldsymbol{X}) =$ 关于 $\boldsymbol{X} = E[(\boldsymbol{X} - \mu_{\boldsymbol{X}})^r]$ 的平均值的 r 阶矩。

习惯上设 $\sigma_{\boldsymbol{X}}^2$ 指随机变量 X 的方差，它是关于平均值的二阶矩 $m_2(\boldsymbol{X})$。一个简单但有用的方程是

$$Var(\boldsymbol{X}) = \sigma_{\boldsymbol{X}}^2 = E[(\boldsymbol{X} - E(\boldsymbol{X}))^2]$$
$$= E(\boldsymbol{X}^2) - (E(\boldsymbol{X}))^2$$

举一个例子有助于澄清几个小问题。设 $\boldsymbol{X}$ 是一个分段连续的随机变量,它的密度 f 的定义是:

$$f(x) = \begin{cases} \frac{1}{2} & \text{如果 } x \in [1, 3] \\ 0 & \text{其他情况} \end{cases} \tag{8}$$

注意,通过(8)定义的密度 f 也能被描述为在样本空间[1, 3]上的均匀密度。事实上,f 在 $x = 1$ 和 $x = 3$ 这两点是不连续的。遵照统计学中的标准实践,我们因而通常在所有实数的集合上,也就是区间$(-\infty, \infty)$,定义密度或分布,在(8)定义的 f 的情况下和函数在某个已知区间之外恒等于零的其他情况下,我们知道,表示函数的曲线下的面积在这个区间之外等于零,因此,我们只需要考虑,在发现矩时,已知区间的积分,等等。这样,对于具有通过(8)给定的密度 f 的随机变量 $\boldsymbol{X}$ 的任何一个函数 h 来说,

$$E(h(\boldsymbol{X})) = \int_{-\infty}^{\infty} h(x) f(x) dx$$
$$= \int_1^3 \frac{h(x)}{2} dx$$

因此,$\boldsymbol{X}$ 的期望值是

$$E(\boldsymbol{X}) = \int_1^3 \frac{x}{2} dx = \frac{x^2}{4}\bigg|_1^3 = \frac{1}{4}(9-1) = 2$$

和 $\boldsymbol{X}$ 的方差是

$$\sigma_{\boldsymbol{X}}^2 = E[(\boldsymbol{X} - 2)^2] = \int_1^3 \frac{(x-2)^2}{2} dx$$

$$= \frac{(x-2)^3}{6}\bigg|_1^3 = \frac{1}{6} - \left(-\frac{1}{6}\right) = \frac{1}{3}$$①

相当令人感兴趣的是，拥有前面考虑的四个经典的离散密度之一的随机变量的平均值和方差。伯努利密度的平均值只等于 p，方差等于 $p(1-p)$。二项密度有平均值 np 和方差 $np(1-p)$。泊松密度是不寻常的，它的平均值和方差是一样的，即都等于 λ。几何密度有平均值 $\frac{1}{p}$ 和方差 $\frac{(1-p)}{p^2}$。

为了明白如何推出这些结果，让我们首先考虑具有几何密度的一个随机变量 $\boldsymbol{X}$。从我们前面对分布和一个随机变量的平均值定义的讨论来看，我们立即像所希望的那样会有

151

$$\begin{aligned} E(\boldsymbol{X}) &= \sum_{k=1}^{\infty} kp(1-p)^{k-1} \\ &= p(1+2q+3q^2+\cdots),\ \text{where } q = 1-p \\ &= \frac{p}{(1-q)^2} \\ &= \frac{1}{p} \end{aligned}$$

我们这里用到的一个熟悉而基本的事实是，求和数列 $1+2q+3q^2+\cdots$。为了求几何密度的方差，我们首先计算二阶原点矩：

$$\begin{aligned} \mu_2'(\boldsymbol{X}) &= \sum_{k=1}^{\infty} k^2 pq^{k-1} \\ &= qp\sum k(k-1)q^{k-2} + p\sum kq^{k-1} \end{aligned}$$

---

① 正是这些简单的计算表明，我们为什么在应用中运用密度比分布要多。很自然，计算规则使用了密度。当然，已知随机变量的密度和分布，都是根据样本空间的概率测度 P 可定义的。

$$= qp \frac{d^2}{dq^2}\sum q^k + p\frac{d}{dq}\sum q^k$$

$$= qp \frac{d^2}{dq^2}(1-q)^{-1} + p\frac{d}{dq}(1-q)^{-1}$$

$$= \frac{2qp}{p^3} + \frac{p}{p^2}$$

$$= \frac{2q}{p^2} + \frac{1}{p}$$

(注意,找到这样一个无穷数列的“窍门”是:这个数列能很容易求和,而且,它的一阶导数或二阶导数是已知的数列。)减去 $E(\boldsymbol{X})^2$,我们就像所希望的那样得到

$$\sigma_{\boldsymbol{X}}^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{q}{p^2}$$

泊松密度的计算更简单: 152

$$E(\boldsymbol{X}) = \sum_{k=0}^{\infty} k\,p(k)$$

$$= \frac{\sum k e^{-\lambda}\lambda^k}{k!}$$

$$= \frac{\lambda \sum e^{-\lambda}\lambda^{k-1}}{(k-1)!}$$

$$= \lambda \sum p(k-1)$$

$$= \lambda$$

对于方差来说,我像所希望的那样有:

$$
\begin{aligned}
\sigma_{\boldsymbol{X}}^{2} &= E(\boldsymbol{X}^{2}) - \lambda^{2} \\
&= \sum_{k=0}^{\infty} k^{2} p(k) - \lambda^{2} \\
&= \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^{k}}{(k-1)!} - \lambda^{2} \\
&= \lambda \sum (k-1) p(k-1) + \lambda \sum p(k-1) - \lambda^{2} \\
&= \lambda^{2} + \lambda - \lambda^{2} \\
&= \lambda
\end{aligned}
$$

为了举例说明这些思想，我们也看一下三个最普通的和最重要的连续或分段连续分布。

第一，在一个区间(a, b)($a < x < b$)的均匀分布，具有密度：

$$
f(x) = \begin{cases} \dfrac{1}{b-a} & \text{对于 } a < x < b, \\ 0 & \text{对于其他情况} \end{cases}
$$

具有这样一个密度的随机变量 $\boldsymbol{X}$ 的平均值是容易计算的：

$$
\begin{aligned}
E(\boldsymbol{X}) &= \int_{a}^{b} \frac{x}{b-a} dx \\
&= \frac{1}{b-a} \left[\frac{x^{2}}{2} \right]_{a}^{b} \\
&= \frac{a+b}{2}
\end{aligned}
$$

153 而且，根据同样的方式我们容易表明：

$$
\sigma_{\boldsymbol{X}}^{2} = \frac{(b-a)^{2}}{12}
$$

第二，根据平均值 μ 和方差 σ^2，在 $-\infty < x < \infty$ 的正态分布或高斯分布有密度，

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

（统计学家称之为正态分布，物理学家称之为高斯分布——相同对象的两个不同名称。）

第三，对于 $-\infty < x < \infty$ 的指数分布有密度

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{对于 } x > 0 \\ 0 & \text{对于其他情况} \end{cases}$$

平均值是 $\frac{1}{\lambda}$，方差是 $\frac{1}{\lambda^2}$。

联合分布。为了分析许多现象，把这些思想推广到几个随机变量是很重要的。为了简化符号，我将只限于连续的或分段连续的变量，但所定义的概念都能直接推广到任意的随机变量。

定义 15. 设 $\mathbf{X}$ 和 $\mathbf{Y}$ 是两个随机变量。

(i) 它们的联合分布是函数 F，使得对于所有的实数 x 和 y：

$$F(x, y) = P(\mathbf{X} \leqslant x, \mathbf{Y} \leqslant y)$$

(ii) 它们的联合密度函数 f 可定义为

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

(iii) 倘若有导数与积分的话，它们的边际密度 $f_1(x)$ 和 $f_2(x)$ 可定义为：

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)dx$$

我们可以表明,这两个随机变量是独立的,如果对于所有的 x 和 y,

$$f(x, y) = f_1(x)f_2(y)$$

在许多方面,对不独立的两个随机变量的测量,理论上最重要的是它们的协方差,我们定义协方差和它们的关联。

定义 16. 设 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 是两个随机变量,它们的平均值和方差是存在的。倘若有重积分的话,$\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的协方差可定义为

$$cov(\boldsymbol{X}, \boldsymbol{Y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - E(\boldsymbol{X}))(y - E(\boldsymbol{Y}))f(x, y)dxdy$$

154 **而且,倘若 $\sigma(\boldsymbol{X}) \neq 0$ 和 $\sigma(\boldsymbol{Y}) \neq 0$,$\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的关联可定义为**

$$\rho(\boldsymbol{X}, \boldsymbol{Y}) = \frac{cov(\boldsymbol{X}, \boldsymbol{Y})}{\sigma(\boldsymbol{X})\sigma(\boldsymbol{Y})}$$

重要的是注意到,如果 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 是独立的,那么协方差和关联就等于零,但逆命题并不总是真的。①

这些思想直接推广到 n 个随机变量 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$。首先,联合分布是函数 F,使得对于所有的实数 $x_1, \cdots, x_n$,

$$F(x_1, \cdots, x_n) = P(\boldsymbol{X}_1 \leqslant x_1, \cdots, \boldsymbol{X}_n \leqslant x_n)$$

第二,已知导数存在,联合密度函数 f 可定义为

---

① 一个物理上有趣而简单的例子是,在量子力学中不会发生相反的情况。这个例子的物理背景是在 7.2 节最后有关 GHZ 型实验所分析的一种量子纠缠。但是,这个例子能够在不参考任何量子现象的条件下完全基于它自己的术语得到理解。设 $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$ 是三个随机变量,它们的值都是 ± 1,设 $E(\boldsymbol{X}) = E(\boldsymbol{Y}) = E(\boldsymbol{Z}) = 0$,再设三阶矩 $E(\boldsymbol{XYZ}) = 1$。于是,很容易证明,协方差等于零,即,$E(\boldsymbol{XY}) = E(\boldsymbol{YZ}) = E(\boldsymbol{XZ}) = 0$。既然三阶矩是确定的依赖性的一个直接陈述,所以,协方差显然是零这一点,并不意味着 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 的独立性。

$$f(x_1, \cdots, x_n) = \frac{\partial^n}{\partial x_1, \cdots, \partial x_n} F(x_1, \cdots, x_n)$$

已知 n 个随机变量的联合密度,它们是**独立的**,当且仅当,

$$f(x_i, \cdots, x_n) = \prod_{i=1}^{n} f_i(x)$$

概率的模态问题。模态问题通常是由概率陈述引起的,我简要地考察一些问题。为了阐明模态思想,再一次考虑把一枚硬币掷三次的简单例子。正如前面详细地讨论过的那样,在三次投掷中,可以看到,样本空间是由与正面和反面的八种可能序列相对应的八种可能结果组成的。我们能够作出关于用样本空间表示的实验的许多模态陈述。特别是,下列陈述,一个是关于可能性的陈述,另一个是关于必然性的陈述,肯定是真的:

(1) 三次试验都得到正面的事件是有可能发生的。

(2) 三次试验,要么至少出现一次正面,要么至少出现一次反面,是必然的。

很容易构造出可能性和必然性的其他陈述,但这样的陈述在概率语境中通常不会太令人感兴趣,因为它们是明摆着的或平凡的。然而,我们通常感兴趣的概率陈述也有一个明显的模态特征。假定在我们的简单试验中,试验是 $p = \frac{1}{2}$ 的伯努利试验。于是,我们把下列陈述也接受为是真的:

(3) 第三次试验得到正面的概率与第二次试验的结果无关。

(4) 已知至少出现两次正面,恰好得到两次正面的概率大于$\frac{1}{2}$。 155

从模态的观点来看,这种方式的语义是相对简单明了的,因为可

能世界的集合恰好是样本空间中可能结果的集合。另一方面，在几乎所有的应用情形中，样本空间中反映出的可能结果的集合受到的限制，比在模态逻辑中通常讨论的可能世界的集合受到的限制要多。许多集合论的关系和其他的数学关系，在样本空间中的可能结果的集合上，被认为是固定的，也是不容易变化的，因为它们通常是对可能世界的记载。

在运用概率概念的数学实践和科学实践中，有一个重要的对模态逻辑的偏离是明显的。概率测度和关于概率的陈述都是以一种纯外延的方式来处理的，而且，概率概念的外延地位与其他科学概念的地位没有差别。在与关于信念的断言相对应的强意向概念似乎发挥自然作用的地方，概率的主观理论也是如此。(有时，某些主观主义者已经考虑过这个问题，但是，它肯定不标准。)

然而，为什么仍然蕴含着概率的模态，有着更深层次的原因。在大多数复杂的应用中，没有提出明显的样本空间，这是事实。实际上，这个问题得到如此的阐述，以至于不存在一个惟一的概率空间，而是有过多的概率空间。一个问题，不管是理论的，还是应用的，都是根据随机变量族来陈述的。一致性的必要条件是，随机变量族有一个与随机变量的所有有穷子集的联合分布相一致的联合分布，但是，对样本空间的定义根本没有惟一性。我所说的，从字面意义上看，似乎违反了前面把随机变量定义为要求它们是在样本空间上定义的函数。然而，在实践中，特殊的样本空间显然完全没有发挥真正的作用，而且，重要的是随机变量的分布，也就是说，随机变量的值的分布。从模态的观点来看，这意味着，可能世界的集合仍然是不明显的。在德·菲内蒂关于概率的两卷本专著中(1974)，他一直不断地努力避免出现随处可见的样本空间。在为了避免混淆术语的地方，他谈到

了随机量(random quantities),而不是随机变量(random variables)。尽管如此,仍需要强调的是,德·菲内蒂的术语和概念分析,与几乎所有的当代统计学家,包括相当同情德·菲内蒂的主观概率观的许多人在内,通常所用的术语与概念分析,并不一致。

概率的不变性。在 4.1 节,我描述了关于不变性的一个重要的逻辑结果,正如塔斯基和吉万特陈述的那样(1987,第 57 页),在全域元素的每一种排列下,一个已知全域的不变的元素之间,只有四种二元关系。这四种关系是通用关系、空关系、恒等关系和相异关系。

这种简单而一般的结果对概率理论不成立,但是,特殊的概率理论的许多特殊的不变性结果是众所周知的。在 4.5 节,描述过一个重要的实例,作为对伯努利和各态历经的马尔可夫过程是完全不变的熵,在下列意义上是同构的:存在着把一个过程映射到保持事件概率的另一个过程。 156

在概率中存在着许多重要的但更基本的不变的结果,我们将把其中的几种结果看成是本节的最终主题。通常有几种不同的方式来阐述同样不变的结果。这里,为了简单起见,我只限于考虑有穷样本空间,即对于一个样本空间来说,可能结果的集合 Ω 是有穷的。正如 4.1 节所讨论的那样,在塔斯基的逻辑结果的精神鼓舞下,更一般地说,在费利克斯·克莱因的几何不变性的爱尔兰根纲领的影响下,第二种限制是,根据把 Ω 一一对应地变换到自身,即,根据 Ω 的双射,阐述这些结果。我只简单地提供对特殊情况的非形式的证明。

独立性。设 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 表示把一枚硬币掷 n 次的 n 个随机变量。这样,Ω 有 2^n 种可能的结果。于是,在 Ω 到自身的任何一种双射的条件下,概率是**不变的**,即是恒定的,当且仅当,随

机变量 $\boldsymbol{X}_i$ 是独立的和公平的，即 $P(\boldsymbol{X}_i = 1$，对于正面来说$)=\frac{1}{2}$。这里是 $n=2$ 的证明。随机变量 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 是独立的，当且仅当，它们的联合密度 $p(x,\ y) = p(x)p(y)$，这里，我们把离散密度写为"p"，取代写为"f"。然后，我们用可能结果的明显符号首先证明

$$p(h_1 h_2) = p(h_1)p(h_2) \tag{9}$$

因为根据我们的不变性假设，这四种可能结果在概率上是相等的，

$$p(h_1 h_2) = \frac{1}{4}$$

但是，$p(h_1) = p(h_1 h_2) + p(h_1 t_2) = \frac{1}{2}$ 和根据相似的论证，$p(h_2) = \frac{1}{2}$。既然 $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$，所以，确立了(9)，而且，对于其他三种可能结果，也能进行相同类型的论证，我们已经表明，硬币是公平的，因为 $p(h_1) = p(h_2) = \frac{1}{2}$。这个证明是平凡的，但它体现了一般方法的起源。

注意，这种情况接近于逻辑结果，而且，理由是显而易见的。根据概率，当我们既有独立性也有公平性时，根本没有区分这两种可能结果的概率方式。它们在概率意义上是等价的。

可交换性。再一次考虑掷硬币的例子，但这一次，我们没有假定公平性。硬币的重心可能是有偏向的。但在关于掷硬币的一般假定的条件下，我们可以假定一种更弱的不变性。我们假定，如果在两种不同的可能结果中，出现正面的次数是相同的，那么，对于有相同正面次数的所有结果来说，Ω 到自身的一种双

射将使概率保持不变。换言之,正如似乎在直观上明确的那样,在 n 次掷硬币的序列中,出现正面的顺序不影响概率。用在某种程度上更适合于 Ω 的双射的术语来说,设已知 n 次试验的结果。于是,在我们的不变性假设的条件下,这种结果的任何一种排列都有相同的概率。例如,考虑把一枚硬币掷三次。然后,用 $\approx$ 表示概率的等价性:

$$hht \approx hth \approx thh,$$

$$htt \approx tht \approx tth$$

但对于 hhh 和 ttt 来说,根本没有这样的等价性。我们对这种可交换的方式也有一种简单的表征可能。也就是说,这种不变性的表征,是通过对随机变量 $\boldsymbol{Y}$ 进行个别计数来捕获到的,在这个例子中,随机变量 $\boldsymbol{Y}$ 的值只是 0、1、2 和 3。这样,$\boldsymbol{Y}(hht)=2$,即出现正面的次数。第 7 节有对可交换性的更多讨论,但这里我们看到,从 2^n 种可能结果戏剧性地简化为只有 $n+1$ 个可能值的一个随机变量 $\boldsymbol{Y}$。当随机变量 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 是可交换的时,所有在概率上能说的,恰好是用随机量 $\boldsymbol{Y}$ 能说的。 157

指数衰变。第三个例子引入了一种不同类型的不变性,这个例子是在第 6 节关于概率的倾向性理论中详细地给出的,但这种不变性在概率理论的许多应用中是相当常见的。我想起了指数分布或它的离散类比(即几何分布)的无记忆特性。如果一个预期的事件直到时间 t 都没有发生,那么它在 Δt 后发生的概率与它在较早时刻 t' 发生的概率没有什么不同。就此问题换一种方式来说,这个过程根本不会老化,从今以后,也不会变得更有可能发生。这里所用的符号是在第 6 节详细讨论这个例子时提出的。设 $p(a_1)$ 是一个镭原子在时刻 t_1 发生衰变的概率,再设已知事件 A_n 在时刻 t_n 不发生衰变,$p(a_{n+1} \mid A_n)$ 是原子在

t_{n+1}试验中发生衰变的条件概率。不变性条件就是：

$$p(a_{n+1} \mid A_n) = p(a_1)$$

设想一下，如果死亡的预期年龄遵守这样一个不变的指数定律，那么生物体的寿命是多么不同啊。当然，有许多物理过程具有这种指数类型。

注意，在后面的两个不变性的例子中，即可交换性和指数衰变，过程的不变特性都没有惟一地确定概率分布，反而对它施加了重要的约束。这是大多数物理过程或心理过程的特征，这些过程决定了密度或分布的形式取决于必须根据经验数据作出评价的少数几个参数。在第 6 节关于概率的倾向性理论中在更多细节上接受了这种观点。

5.2 概率的经典定义

考虑到普遍存在的概率特征和人类有史以来，即现在和贯穿于大部分的历史记载，对赌博游戏的广泛兴趣，发现关于概率的系统讨论的智力发展是很晚的，这是令人惊讶的。除了 16 世纪的少数几部著作之外，在 17 世纪之前根本没有重要的概率理论的作品。尽管在希腊数学和天文学中有深入的发展，但是，在希腊的思想中，显然没有系统地研究过概率问题。在希腊科学或数学中似乎根本没有作为概率现象讨论的处理测量误差的问题，即属于定量分析的问题。

158 在传统意义上，系统的概率理论的主要起源归功于 17 世纪的帕斯卡尔(Pascal)和费马(Fermat)的工作，而且，现代的证据似乎很少对这种归因提出挑战。传统的解释是，舍瓦利耶·德·梅雷(Chevalier de Méré)向帕斯卡尔提出了包括计算赌博

游戏中输赢的可能性在内的一些问题。帕斯卡尔与费马就这些问题进行了通信讨论,通过这种通信形成了概率的数学理论。令人感兴趣的是注意到,在帕斯卡尔和费马的通信中现存的第一封信里,即日期是 1654 年 7 月 29 日的信里,所讨论的问题已经是相当难的一类问题。被称为**得分问题**。两位赌徒在赌博时,每一位都需要一个获胜的已知分数,每一步都有相等的机会得 1 分。假定赌徒在没有结束赌博的条件下分开。问题是确定应该如何根据他们当前的得分划分赌金。这与询问在任意特定阶段每位赌徒在赌博中获胜的概率一样。帕斯卡尔没有完全解决这个问题,我们这里也不讨论这个问题。他确实提供了在得 6 分的赌博中可能发生的所有情况,他也涵盖了当一位赌徒只差 1 分就能获胜和第二位赌徒没有得分时的情况,还有当一位赌徒得了 1 分和第二位赌徒没有得分的情况。

在给费马的同一封信中,帕斯卡尔也讨论了在计算把一枚骰子掷四次能否掷出一个 6 点的可能性(这种可能性是 671 比 625)和用两枚骰子能否掷出两个 6 点的可能性时所产生的一个悖论问题。掷出两个 6 点的可能性并不支持掷二十四次时的可能性,尽管事实是,在 36 是掷两枚骰子时出现的情形数和 6 是掷一枚骰子时出现的情形数的情况下,24 比 36 相当于 4 比 6。这个问题是一个简单的问题,而且,帕斯卡尔也轻松地处理了它,即使德·梅雷声称,他在计算时发现了这个明显悖论的矛盾根源。我这里之所以提到这个例子,是因为我想在后面强调,过去在一开始讨论概率问题时,大概由于一枚骰子的形状是明显对称的,所以,已经轻易地和自然地运用了这样的经典定义:一个事件的概率是有利的情形数除以所有可能的情形数。

对 17 世纪和 18 世纪的概率理论的发展史感兴趣的读者,可参考托德亨特(Todhunter)的标准著作(1865,1949 年重印),

还有更新近的哈金(Hacking)的较少数学和较多哲学的解释(1975)。尽管古人早对赌博发生了很大的兴趣,但是,关于早期赌博史的讨论,包括关于概率理论为什么没有被更早地提出的某些有趣的推测,可以在大卫(David)的著作中找到(1962)。除此之外,在桑伯斯基(Sambursky)的一篇论文中包含了对古希腊思想中的可能概念和或然概念的详细讨论。

在帕斯卡尔和费马正在著书立说的 17 世纪中叶和拉普拉斯成为统治性人物的 18 世纪之间的这段时期,关于概率的最重要的著作可能是:(i) 惠更斯首版于 1657 年的一本早期论著《论赌博中的计算》(*De Ratiociniis in Ludo Aleae*),(ii) 伯努利(Jacob Bernoulli)首版于 1713 年的《猜测术》(*Ars Conjectandi*),以及(iii) 棣美弗(de Moivre)首版于 1718 年的159 《机会论》(*The Doctrine of Chances*)。在这个时期,拉格朗日、欧拉(Euler)等人获得了重要的数学结果。此外,受人尊敬的托马斯·贝叶斯的著名论文身后在皇家学会的哲学学报第 1763 卷发表。

拉普拉斯。然而,拉普拉斯的工作达到了这个时期的顶峰,他的工作开始于 1774(a,b)年的两本很厚的论文集,并在 1812 年首次出版的论著《概率的分析理论》(*Théorie Analytique des Probabilités*)中达到了最高点。这本大部头的著作在概念上和数学上都是很深刻的和成熟的。从我们这里感兴趣的观点来看,重要部分是导言中对概率基础的概念分析。这个导言曾在 19 世纪初单独发表,所用的英语译文的标题为"关于概率的哲学论文"(*A Philosophical Essay on Probabilities*)(1951)。因为拉普拉斯的这篇论文是概率的经典定义的标准来源,所以,我们想分析拉普拉斯观点的核心内容。

首先,重要的是强调,对于拉普拉斯来说,宇宙是决定论的,

概率只产生于我们对事件的确切原因的无知。他在第2章的开头几行很简洁地提出了这个问题：

> 所有的事件，即使是似乎不遵守伟大的自然定律因而是无意义的那些事件，其结果都恰好与太阳的旋转一样是必然的。在对把这样的事件与整个宇宙系统联合起来的纽带一无所知的情况下，按照它们发生并被有规则地重复或无序地出现，使得它们依赖于终极原因或依赖于机遇；但这些想象的原因随着知识面的扩展逐渐地退出，而且，在可靠的哲学面前完全消失了，这表明，这些想象的原因只表达了我们对真正原因的无知。
>
> 一个事件，如果没有产生它的原因，就不可能发生，基于这个明确的原理，有一条纽带把当前的事件与过去的事件联系起来。这个公理，被称为充足理由律，甚至可推广到平常考虑的行为。
>
> (Laplace, *A Philosophical Essay on Probabilities*: p. 3)

拉普拉斯在这一段之后写下了他关于宇宙的决定论特征的著名陈述。从当前了解到的宇宙状态来看，一位"智者"，或者，也许像我们用现代术语所说的那样，"一台适当地计算数据的计算机"，能够确定整个宇宙的过去和未来。

> 我们就应该把宇宙当前的态看成是它过去态的结果和未来态的原因。已知某一瞬间有一位"智者"，他能够理解使自然界充满活力的所有的力，也能理解构成自然界的生命的各种情境——分析这些巨大数据的一位智者——它信奉，宇宙中最大天体的运动和最轻原子的运动遵循相同的公式；对它来说，没有什么是不确定的，而且，未来像过去一

样会呈现在它的眼前。它能使天文学尽善尽美,在这方面,人类思想提供的想法,对于这位智者来说,是微不足道的。160 人类在力学和几何学中的发现再加上万有引力的发现,使其以相同的分析表达理解世界系统的过去和未来的态。人类把相同的方法应用到所认识的其他对象,已经成功地谈到所观察现象的一般定律并预言了特定情况应该产生的那些定律。寻找真理的所有这些努力,往往会不断地把人类领向我们刚才提到的那位分析巨大数据的智者,但是,这是永远达不到的。人种特有的这种趋势是,使它比动物高级;他们在这方面的进步区分出国家和时代,并且构成了他们真正的荣耀。①

(Laplace, *A Philosophical Essay on Probabilities*: pp. 4 - 5)

当我们转向讨论概率定义本身的形式陈述时,有必要意识到,对于拉普拉斯来说,概率具有与无知的这种主观联系。在这种意义上,概率的经典定义与主观的概率观非常一致,与下一节讨论的客观的相对频率的那类观点相反。

在第3章,拉普拉斯转向了概率计算的一般原理。他直接从头两个原理的陈述开始,其中的第一原理恰好是概率的经典定义。

第一原理——这些原理的第一原理是概率的定义本身,正如已经看到的那样,它是有利情形的总数与所有可能

① 罗杰·哈恩(Roger Hahn)提供的重要的历史观点是,拉普拉斯在他的一本早期论文集中,实际上陈述了他关于决定论的观点(1773)。正如哈恩所表明的那样,他显然受到了由孔多赛(Condorcet)给达朗贝尔(D'Alembert)的一封公开信中(1768)提出的更早的类似陈述的影响。

情形的总数之比。

第二原理——但假定各种不同情形都是同样可能的。如果它们并非如此,我们将首先确定它们各自的可能性,对可能性的精确评价是机遇理论的最微妙的观点之一。于是,概率是每一种有利情形的可能性的总和。

(Laplace, *A Philosophical Essay on Probabilities*: p. 11)

拉普拉斯继续陈述了我随后要讨论的许多加法原理;但刚才引用的头两个原理构成了概率的经典定义的核心,因此,在这里对它们的含义进行详细考察是合乎情理的。在转向分析拉普拉斯的定义时,我应该一开始就强调指出,我不再关注概率理论的历史,尤其不关注对经典理论的批评,因为这些批评是在历史上逐步提出来的。相反,我想集中于对这个定义及其不足的逻辑分析。

从形式的观点来看,很容易明白,如何能够推广第1节的定义2,使得概率的形式定义与拉普拉斯的第一原理所阐述的经典定义相一致。为了简单起见,让我们只限于可能结果的有穷集合的情形。因为拉普拉斯的第一原理包括的内容,使定义2的多数公理成为多余的,所以,也许有益的是,首先重新开始陈述一个定义,然后断言把这个定义的模型与定义2的模型联系起来的一个简单定理。在这个新定义的陈述中,我们用了康托(Cantor)的双杠符号表示一个集合的基数;例如,$\overline{\overline{A}}$是集合A的基数。

定义 1. 一个结构$\Omega=(\Omega, \mathfrak{F}, P)$是一个有穷的拉普拉斯概率空间,当且仅当, 161

1. Ω是一个有穷集;
2. $\mathfrak{F}$是Ω上的一个集合代数;

3. 对于在 $\mathfrak{J}$ 中的 A，

$$P(A)=\frac{\overline{\overline{A}}}{\overline{\overline{\Omega}}}$$

似乎明显的是，定义 1 以一种明确的形式方式表达了拉普拉斯的第一原理的内容。下列定理也显然是这个定义的一个平凡结果。

定理 1. 任何一个有穷的拉普拉斯概率空间$(\Omega, \mathfrak{J}, P)$，在第一节的定义 2 的意义上，都是一个有穷地相加的概率空间。

对于赌博游戏的讨论，当然是在 17 和 18 世纪最流行的概率问题，拉普拉斯的概率空间似乎是自然的和适当的，因为强加于骰子的对称性或随机地洗一副纸牌，很容易致使满足定义 1 的公理 3。

在我们借助特殊的理由，拒绝把拉普拉斯的概率空间看成是阐述所有概率问题的适当方法之前，我们应该考虑我们所用的论证中所包括的一个重要的哲学问题。这个问题的本性可以通过与一个自然数的集合论构造的对比来搞清楚。在数学的集合论发展中，现在，习惯于把自然数等同于有穷基数或有穷序数。事实上，如果我们根据冯·诺伊曼构造的序数，我们就可以有自然数、有穷基数和有穷序数三方的等同关系。这里，除了强调任何一种此类构造的检验标准之一都满足皮亚诺(Peano, Giuseppe)的自然数公理之外，给出如何在集合的一般公理提供的框架内获得这种构造的技术细节，是不适当的。现在，一个熟悉的数学史的事实是，在阐述这些公理之前，或者，在实现自然数的任何序列的集合论构造之前，关于自然数的困难的和精巧的定理已经得到了证明。像皮亚诺的公理系统或像冯·诺伊曼的集合论构造，都没有给出这些定理，它们遭到了拒绝，因为完全不依赖于这些公理或构造，我们有一个明确概念判断自然数

的真伪，或更保守地说，我们有一个明确的概念愿意把自然数接受为是真的或是假的。在我看来，对这种明晰性的理由的适当解释，是根据基于算术基础的经验事实给出的，但是，支持算术最初是一门经验科学的理由，不可能在这项工作中得到探究。在关于概率基础的一系列立场的讨论中，不得不面对的是，我们将基于很难定义的直觉，提出是接受还是拒绝各种概率空间的论证。在大多数情况下，我以为，这个例子或反例的适当性的直觉基础是足够明确的，不需要支持它的任何一般的方法论论证。我们拒绝接受这种自然数的任何一个公理系统：它们对偶数和奇数的解释与自毕达哥拉斯(Pythagoras)以来公认的熟悉的真理不相符；同样，如果对有偏向的骰子或不公平的硬币的实验没有提供一种直观上可接受的描述，那么我们拒绝接受应该如何构造概率空间的任何一种解释。一旦提出一个明显的反例，这样的构造为什么是不可接受的理由，我就认为是十分明显的。 162

当我们转向对概率论的经典定义进行详细的分析时，一开始就值得注意的是，很难使拉普拉斯的第一原理和第二原理相互一致起来。第一原理提供了一个不含糊的、明确的、但稍加限制的概率定义，本节前面的定义 1 的公理 3 刚好体现了这一点。另一方面，第二原理对没有同等可能性的情形所提供的陈述，是相当模糊的。我们暂且忽略第二原理，只强调第一原理。当我们试图坚持第一原理中包含的定义时，关于表征有偏向机制的实验立即产生困难。例如，在经典定义中，我们如何表示把一个有偏向的硬币掷两次的可能结果呢？就明晰性来说，假定我们正在基于大量的实验来处理一枚硬币，我们发现，在标准的投掷过程中，出现正面的概率是 0.55。显然，在以一种自然方式表示把这样的一枚硬币掷两次的结果的定义 1 的意义上，无法定义一个拉普拉斯的概率空间；因为事件结构以一种明确的方式

确定了集合 Ω 应该是什么，而且，一旦确定了这种事件结构，公理 3 就没有什么可补充的。按照定义 1 的公理 3，对确定概率的事件结构的这种观察，是关于拉普拉斯概率空间在概念上重要的而在数学上无关紧要的结论。从定义 1 来看，应该明显的是，拉普拉斯的概率概念，根据可能结果的集合 Ω 和事件代数 $\mathfrak{J}$ 是可定义的。这意味着，从概念的观点来看，在超越 $\mathfrak{J}$ 是 Ω 上的一个集合代数(或 σ-代数)的必要条件，一般地阐述拉普拉斯的概率空间时，根本没有重要的假设。此外，在有穷情形中自然如此，我们简单地把 $\mathfrak{J}$ 看成是 Ω 的所有子集的集合，而且，除了可能性的集合有可能是有穷的之外，我们没有任何类型的重要公理。①

现在，可能有人认为，应该用掷骰子或纸牌游戏中所用的对称性原理，取代第 1 节定义 2 的明显重要的公理，但很清楚，从形式的观点来看，这些对称性原理是多余的。我的意思是说，事件结构惟一地确定了样本空间 Ω。如果我们认真地接受经典定义，那么，无论是在把一枚骰子掷三次的例子中，还是在把一枚硬币掷两次的例子中，样本空间 Ω 的定义都是通过描述可能结果根本不是对称性的任何考虑惟一地确定的。对称性原理只能从直观上告诉我们，经典定义的特殊应用是否是明智的。为了把对称性的任何原理合并到概率的形式定义中，以下列方式弱化经典定义是有必要的。我们需要使应用拉普拉斯第一原理依赖于满足几何对称性和物理对称性的某些原理条件。从一种观点来看，这似乎是一个要遵守的合理程序，因为在所有熟悉的应用情形中，显然某些这样的原理得到了满足，而且事实上，是直

163

① 第 5 节的定理 3 表明，一旦引入无穷小的方法，从一种形式的观点来看，如何能把拉普拉斯的概率空间推广到提供一个几乎相等的任意概率空间。

观上可接受的应用经典定义的关键。另一方面,把这样的原理明显地引入概率的系统定义中,就会使形式的任务非常复杂,因为根本不清楚人们能够多么满意地陈述所包括的一般的对称性原理。①

经典悖论。我现在转向考虑,当我们试图应用定义1时,所产生的几个经典悖论。

伯川德悖论。考虑下面的简单问题。在一个瓶子里倒入水和酒的混合液。只知道在这种混合液中,水与酒的比例是不小于1和不大于2。在以这种明显的方式把定义1的公理3推广到连续的情形时,公理3中表示的对称性原理或无差别原理(the principle of indifference)指出,我们应该假定,当水与酒之比在1和2之间时,位于这个区间的任何一个地方的混合液的概率,是由这个区间的一种均匀分布产生的。因此,位于1和1.25之间的混合液的浓度的概率是0.25。位于1.6和2之间的混合液的浓度的概率是0.4,等等。所有这些似乎是足够简单明了的。

然而,现在考察看问题的另外一种方式。我们考虑的不是水与酒的比例,而是水与酒的反比。从最初陈述问题时提供的信息来看,我们知道,酒与水之比位于1/2和1之间。我们再一次假定在这个区间的一种均匀分布。这样,在1/2到3/4的区间内,酒与水的实际之比的概率是0.50。现在,3/4的酒比水对应于4/3的水比酒,但是,根据均匀分布假设,从提出比率的这两种方式,我们却得出一种矛盾。设“H”代表水,“W”代表酒。对于

---

① 相反,特殊的对称性原理很容易在应用中陈述和使用。例如,参见第1节最后关于不变性的讨论,或第6节最后关于定性的离散密度的讨论。

$$P\left(\frac{4}{3} \leqslant \frac{W}{H} \leqslant 2\right) = \frac{2}{3} \tag{1}$$

和

$$P\left(\frac{1}{2} \leqslant \frac{H}{W} \leqslant \frac{3}{4}\right) = \frac{1}{2} \tag{2}$$

但是，

$$P\left(\frac{4}{3} \leqslant \frac{W}{H} \leqslant 2\right) = P\left(\frac{1}{2} \leqslant \frac{H}{W} \leqslant \frac{3}{4}\right) \tag{3}$$

因为数学事实是，

$$\frac{4}{3} \leqslant \frac{W}{H} \leqslant 2 \text{ 当且仅当}, \frac{1}{2} \leqslant \frac{H}{W} \leqslant \frac{3}{4}$$

显然，方程(1)、(2)和(3)是相互矛盾的。彭加勒(Poincaré)
164 把这个悖论命名为“伯川德悖论”(Bertrand Paradox)，他是根据第一次讨论这个例子的一位法国数学家的名字命名的(Bertrand, 1888)。

伯川德悖论中的这种困难源于引入了一个这样的随机变量：它的值域是一个无穷集，在这个案例中，是实数的一个有穷区间。不是通过计算这个随机变量的可能值的总数和把先验概率$\frac{1}{N}$分配给每个点，而是通过测量这个区间的长度和把相等的先验测量分配给任意两个等长的子区间来确定均匀分布。于是产生了困难，因为我们可以考虑这个随机变量的一种变换，这种变换不是一种线性变换，在这个案例中，是从一个数过渡到它的倒数的非线性变换。有人可能认为，我们仅限于一个有穷的点数就能避免这个悖论；但是，一旦我们处于像现在这样在点集上有一种自然度量或测度的一种情形，就一定并非如此。例如，如

果我们在考虑水与酒的比例时限于有穷点集

1.1, 1.2, 1.3, 1.4, 1.45, 1.46, 1.47, 1.5, 1.8, 1.9,

那么,令大多数人确实感到震惊的是,发现我们把相等的权重分配给所有这些点,特别是,如果我们从一般的无知假设开始,理解比例不在指定的 1 和 2 之间的混合液的本性。这种惊奇的理由是显然的;即在可能值的区间上所选择的点,在我们采用的自然度量中,不是等间隔的。如果我们认真地采纳这样一种自然度量,那么,当然,即使是对于有穷点集来说,也会产生这种悖论。例如,如果我们在下列集合中选择等间隔的比率

1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0,

很自然的是,把一个相等的先验概率分配给每个比率,但当我们考虑反问题并看看酒与水的比率时,这个相同的集合就成为如下形式

0.50, 0.53, 0.56, 0.59, 0.63, 0.67, 0.71, 0.77, 0.83, 0.91, 1.00,

而且,把一个相等的先验概率分配给这种不是等间隔的数集的项,似乎是很不自然的。下面是一个更加古老的问题。

布封投针问题(Buffon's needle problem 1733,1777)。在一个地板上划出许多条平行线,然后,随机地投下一根针。这根针与地板上的一条线相交的概率是多大呢?我们关心的是,预言这根针相对于地板上的平行线系统的位置。现在,针的位置能用坐标系来描述,那么矛盾来自这样的事实:可能使用通过非线性变换相关的不同类型的坐标系。两个熟悉的例子是常见的笛卡儿坐标和极坐标。当然,如果我们指定一种均匀分布,首先是相对于笛卡儿坐标的,然后是相对于极坐标的,那么我们肯定会获得不同的结果。

伯川德的随机弦悖论(Bertrand's random chord paradox 1888)。这个问题是,在一个单位半径的圆内,选择一条长度大

于$\sqrt{3}$的“随机弦”的概率是多大(在这里,当然$\sqrt{3}$是一个内切等边三角形的边长)。正如伯川德所评论的那样,这个问题有三个显明而简单的解。

165

1. 因为这个圆的任何一条弦与这个圆相交于两点,所以,我们可以假设,这两点独立地分布在这个圆的圆周线上。如果不失一般性的话,我们可以假设,两点中有一点是内切三角形的顶点。等边三角形的三个顶点就把这个圆周划分为三个相等的部分,这样,要使弦大于$\sqrt{3}$,另一个点恰好落在 1/3 的圆周上。因此,找到这样一条弦的概率是 1/3。

2. 如果我们现在直接考虑这条弦的长度,我们会立即看到,弦长取决于它离圆心的距离且与其方向无关。为了表示弦离开圆心的距离这个事实,我们可以假设,把弦的方向固定在垂直于圆的一条已知直径的方向上,用沿这条直径的距离来测量弦离开圆心的距离。我们然后可以假设,弦与这条直径的交点有一种均匀分布。对于一条长度大于$\sqrt{3}$的弦来说,交点离圆心的距离一定小于 1/2,这样,找到这样一条随机弦的概率是 1/2。

3. 任何一条弦都是通过弦与垂直于弦的一条半径的交点惟一地定义的。一个自然的假设是,这个交点均匀地分布在这个圆上。于是,弦位于面积 A 的任何一个区域的概率是 A/π,因为这个圆的总面积是 π,因此,获得长度大于$\sqrt{3}$的一条随机弦的概率是 1/4。

毫无疑问,也能用弦的其他自然特性产生除这三种情况之外的分布。① 概率的经典定义的应用不是以概念上明确而清楚

① 注意,关于这一点是不明显的,我们已经转向潜在地使用了具有一个结果的无穷集的概率空间。在 20 世纪之前,对这些案例没有作出明显的形式分析,而且,在柯尔莫哥洛夫的经典工作(1933)提供了一条广泛承认的进路之前也很少见,这体现在第 1 节作为定义 2 给出的概率空间的定义中。

的方式来决定的，这个例子应该使这一点变得明确起来。重要的是注意到，在选择了一个特殊的概率空间之后，就不会产生这个问题。这一节的定义 1 的应用，特别是定义 1 的公理 3 的应用，对于一个特殊的概率空间和在这个空间上的集合代数来说，是十分简单明了的。既不明确也不明显的是，应该选择多么特殊的概率空间，才能以适当方式描述一个给定的问题。我们再一次必须超越定义 1，要求应该用更基本的对称性原理来构造适当的概率空间。就我所知，尽管在第 5 节讨论的确证理论中对特殊测量的选择，真的取决于某些简单的对称性原理的应用，但是，一直没有对这些对称性原理进行过广泛的讨论。另一方面，在确证理论中关于对称性原理的讨论，还没有超越最初的阶段，而且，不足以确定，把什么样的先验概率测度运用于一个简单但相对精致的像随机弦之类的问题中。对寻找在所有情况下都承认的任何一个简单而自然的对称性原理的绝望也许是捍卫概率的主观理论的主要思路之一，第 7 节讨论这个问题。 166

关于拉普拉斯原理 3－10 的历史注释。拉普拉斯关于概率的前两个原理阐述了概率的经典定义，与这两个原理相联系，我提到了也对拉普拉斯陈述的其他原理给予关注。我认为，一个富有历史兴趣的问题是，对这些原理进行简要的评论，并理解它们如何与本节给出的经典定义的形式重构相关。拉普拉斯在我们已经谈到的《关于概率的哲学论文》的第三章和第四章陈述了这些额外的原理。

第三原理只陈述了独立事件的乘法规则。第四原理陈述了非独立事件的乘法规则，因而，像在第 1 节中给出的那样，只是条件概率定义的一种形式。第四原理的形式就是

$$P(A \cap B) = P(A \mid B)P(B)$$

第五原理恰好是对第四原理的另一种表述,在这种情况下,形式上我们用作条件概率的定义是

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

第六原理是对贝叶斯定理的一种陈述。也许重要的是注意到,在第六原理的表述中,拉普拉斯明显地考虑了,倘若原因或假设的先验概率可能是不相等的,如何表述这个定理的问题。因此,他的表述正好对应于第1节作为定理8给出的贝叶斯定理的表述。第七原理是关于作为定理7给出的总概率的定理的一种表述。

他从第八原理开始介绍第四章。由于要考虑后三个原理,我们从纯概率的问题转向期望值问题,或者,像拉普拉斯提出的那样,关于希望的问题。第八原理告诉我们,如何从引起一个有利结果的事件发生的概率来计算数学期望值。拉普拉斯的简单例子将使这一点变得更清楚:

> 让我们假定,在正反面的游戏中,如果保罗第一次掷出正面时,得到两法郎,只在第二次掷出正面时,得到五法郎。两法郎乘以第一种情况的概率 1/2 和五法郎乘以第二种情况的概率 1/4,乘积之和,即 2 又 1/4 法郎,将是保罗的优势。给出他的这种优势的那个人,应该预先给出这个和;例如,为了保持游戏的平等性,应该相等地掷出它产生的优势。
>
> 如果保罗第一次掷出正面,得到两法郎,第二次掷出正面,得到五法郎,不管他第一次投掷是否能得到正面,第二次投掷时,得到正面的概率是 1/2,1/2 乘两法郎加五法郎,得到的乘积之和是 3 又 1/2 法郎,这是保罗的优势,因而也

是他在这个游戏中的风险。

(Laplace, *Philosophical Essay on Probabilities*: pp. 20 - 21)

第九原理把第八原理推广到能够产生输赢的情况。最后,第十原理是对丹尼尔·伯努利(Daniel Bernoulli)的道德期望原理(principle of moral expectation)(1738)的不太明确的语言表述,本章第7节以预期效用的名义讨论这个原理。 167

原理3-8都是第1节给出的概率空间的形式定义的相当直接的结果,事实上,并不依赖于概率本身的经典定义,这应该是很清楚的。

最后,重要的是评论,即使概率的经典定义是模糊的和难以应用的,但它仍然是重要的,因为实际上,在对称性的指导原理是明确的和确实是普遍认同的许多重要应用中,正是用了概率的这个定义。在赌博游戏中,正是用概率的经典定义进行标准的计算。

5.3 无穷随机序列的相对频率理论

这一节致力于讨论相对频率理论(例如,参见Nagel 1939b)。我们首先从一个定理开始。在这之后是举例和批评,这样做的宗旨是,考察在应用概率理论时使用定理的合理性。这个定理用了长度为N的有穷序列概念。这样一个序列正好是其定义域前N个正整数的一个函数。这个定理的形式陈述应该没有掩盖它恰好用相对频率定义概率的简单思想。

定理1. *设f是长度为N的一个有穷序列,设$\mathfrak{F}$是在f的值域上的一个集合代数,也就是说,集合$R(f)$,再设对于在$\mathfrak{F}$*

中的每个 A,函数 P 可定义如下:

$$P(A) = \frac{\overline{\overline{\{i: f_i \in A\}}}}{N}$$

(即,使 f_i 在 A 中的 i 的个数除以 N 等于 $P(A)$ 的值)。于是,$(R(f), \mathfrak{J}, P)$ 是一个有穷的相加概率空间。

证明: 显然,$P(A) \geqslant 0$。既然 $\overline{\overline{\{i: f_i \in R(f)\}}} = N$,那么也有 $P(R(f)) = 1$。但如果 A 和 B 在 $\mathfrak{J}$ 中,而且 $A \cap B = \varnothing$,那么

$$\begin{aligned} P(A \cup B) &= \frac{\overline{\overline{\{i: f_i \in A \cup B\}}}}{N} \\ &= \frac{\overline{\overline{\{i: f_i \in A\}}}}{N} + \frac{\overline{\overline{\{i: f_i \in B\}}}}{N} \\ &= P(A) + P(B). \end{aligned}$$

作为例示上述定理的一个例子,设 f 是投掷硬币二十次的记录。因此,f 记录了一个实验。重要的是注意到,f 不是二十次投掷的所有可能结果的集合,而是某种特殊的结果。根据一个很简单的相对频率理论,我们用这个实际结果计算概率。记住我们现在研究应用问题,并希望知道如何分配实际概率。已知

168

$$f = thttthhththhthhttht$$

假定 A 是得到正面的事件,那么

$$P(A) = \frac{\overline{\overline{\{i: f_i \in A\}}}}{N} = \frac{9}{20}$$

我们将对根据经验定义概率的上述方法的主要批评——通过用一个有穷试验序列——加以讨论。随后提出可替代的经验

方法。

相对频率理论意在成为一个客观理论。给定事件的概率独立于我们关于该事件的知识。换句话说,对于相对频率的理论家来说,一个已知硬币在一次投掷中出现正面的概率,是硬币和掷硬币的实验条件的一个物理属性;它不是我们是否有该硬币的或投掷的知识的一个函数(正如一位主观主义者所说的那样)。

从形式的观点来看,应用概率的相对频率理论的问题,最初是确定哪些有穷相加概率空间是可接受的问题。相对频率观点的基本思想是,应该以定理 1 例示的方式从(有穷的或无穷的)序列产生概率空间。难题一直是为这个**随机**序列的概念找到一个适当的定义。我们在本节将考虑几种可能性,最后以丘奇的定义结束,但首先,我希望提到用一个有穷序列作为基础所受到的批评。

对有穷序列的两个主要批评是:(i) 这样定义的概率是不稳定的;(ii) 很难定义有穷序列的随机性。例如,在上述掷硬币的例子中,如果在已知的二十次试验中增加新的试验,那么 $P(A)$可能发生很大的变化。当然,正好二十次试验的选择是任意的。坚持进行大量的试验,至少能够部分地应对关于不稳定性的批评。但是,即使试验很多次,也很难给随机性下一个适当的形式定义。因此,从数学的观点来看,自然的问题是"超越这种限制",考虑无穷序列。后面将谈到无穷序列的这种引入受到的经验批评,在下一节,我们返回来详细地考虑有穷随机序列的问题。

在经典意义上,无穷序列一直是相对频率理论的基础。与无穷序列相关所产生的形式问题,已经引起了很多数学家的关注。在某些数学预备知识之后,考察了随机序列的三个定义,每

个后面的定义都比其前面的定义更强。

一个(无穷的)序列是一个函数,它的定义域是正整数集合。如果 s 是一个序列,那么 s_n 是序列 s 的第 n 项。一个正整数序列是它的所有项都是正整数的一个序列。

设 s 是一个实数序列,那么 $\lim n\rightarrow\infty\ s_n = k$, 当且仅当,对于每个 $\epsilon > 0$, 都有一个整数 N,使得对于所有的整数 n 来说,如果 $n > N$, 那么

$$|s_n - k| < \epsilon$$

定义 1. 如果 s 和 t 是实数序列,那么

169 (i) $s+t$ 是使得 $(s+t)_n = s_n + t_n$ 的序列,

(ii) $s \cdot t$ 是使得 $(s \cdot t)_n = s_n \cdot t_n$ 的序列,

(iii) 如果对于每个 n, $s_n \neq 0$,那么 $\frac{1}{s}$ 是使得 $\left(\frac{1}{s}\right)_n = \frac{1}{s_n}$ 的序列。

略去关于极限的下列基本定理的证明。

定理 2. 如果 s 和 t 是有极限存在的序列,那么

(i) $\lim(s+t) = \lim s + \lim t$,

(ii) $\lim(s \cdot t) = \lim s \cdot \lim t$,

(iii) 如果对于每个 n, $s_n \neq 0$, 那么,如果 $\lim s \neq 0$, 则 $\lim \frac{1}{s} = \frac{1}{\lim s}$。

我们下面引入一个频率函数。

定义 2. 设 s 是一个序列,再设 $\mathfrak{J}$ 是 $R(s)$(s 的值域)的所有子集族。在 I^+ 是所有正整数的集合的地方,设 t 是在 $\mathfrak{J} \times I^+$ 上定义的函数,使得对于在 $\mathfrak{J}$ 中的所有 A,

$$t(A, n) = \overline{\overline{\{i: i \leqslant n \ \&\ s_i \in A\}}}$$

我们把数$\frac{t(A, n)}{n}$称为A在s的前n项中的**相对频率**。如果当$n \to \infty$时，存在有函数$\frac{t(A, n)}{n}$的极限，那么这个极限是A在s中的**极限相对频率**。

定理 3. 设s是一个序列，$\mathfrak{F}$是在$R(s)$上的一个集合代数，使得如果$A \in \mathfrak{F}$，那么A在s中的极限相对频率存在。设P是在$\mathfrak{F}$上定义的函数，使得对于$\mathfrak{F}$中的每个A，

$$P(A) = \lim_{n \to \infty} \frac{t(A, n)}{n}$$

那么$R(s)$，$\mathfrak{F}$，$P)$是一个有穷相加概率空间。

证明：既然对于每个n，$\frac{t(A, n)}{n} \geqslant 0$，那么肯定

$$\lim_{n \to \infty} \frac{t(A, n)}{n} \geqslant 0$$

此外，对于每个n，$t(R(s), n) = n$，因此，

$$\lim_{n \to \infty} \frac{t(R(s), n)}{n} = 1$$

最后，如果$A \cap B = \varnothing$，那么

$$\lim \frac{t(A \cup B, n)}{n} = \lim \left[\frac{t(A, n)}{n} + \frac{t(B, n)}{n}\right]$$
$$= \lim \frac{t(A, n)}{n} + \lim \frac{t(B, n)}{n}$$

因此，我们看到，$(R(s)$，$\mathfrak{F}$，$P)$是一个有穷相加概率空间。

上述无穷序列表述的一个例子如下：考虑掷硬币的序列

$$(h, h, t, h, h, t, h, h, t, h, h, t, \cdots, h, h, t, \cdots)$$

显然，如果 A 是由出现的正面所组成的事件，也就是说，如果 $A=\{h\}$，那么

$$\lim\left[\frac{t(A,\ n)}{n}\right]=\frac{2}{3}=P(A)$$

170 另一方面，如果 $B=\{t\}$，那么

$$\lim\left[\frac{t(B,\ n)}{n}\right]=\frac{1}{3}=P(B)$$

尽管随机性的直观概念完全是不明确的，但似乎显而易见的是，这个投掷序列不是随机的。基本单元 hht 的重复，诱使我们为此寻找某种因果性说明。如果概率是这个系统的一个客观特征，那么任何一个简单选择的序列所产生的事件概率，都应该与最初的无穷序列产生的事件概率是相同的。更确切地说，既然一个结果的概率是这个序列的一个客观属性，所以，如果我们选择的任何一个无穷子序列都具有这样的特性：选择出的子序列的每一项都独立于这一项的结果，那么我们将期望得到相同的数值概率。但在我们的例子中，如果我们得到的子序列是，从第三项开始并由其后的每个第三项构成，那么我们得到无穷序列 $(t,\ t,\ t,\ t,\ t,\ \cdots,\ t,\ \cdots)$，在这个序列中，得到反面的概率是 1。

极限相对频率并非总是存在的，下列例子能够表明这一点。考虑由下列方式形成的只包括 1 和 0 的序列：

$$s=(1,\ 0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 1,\ \cdots)$$

在这里，如果 k 是使得 $2^{k-1}\leqslant n<2^k$ 的偶数，那么 $s_n=0$，如果 k 是使得 $2^{k-1}\leqslant n<2^k$ 的奇数，那么 $s_n=1$。如果 a 是偶数，那么在前 2^a-1 项中出现 1 的相对频率是 1/3。如果 a 是奇数，那么在前 2^a-1 项中出现 1 的相对频率大于 2/3，但当 a 增加时，

逐渐地逼近 2/3。

此外，可能有一个序列 s 和一个 $R(s)$ 的子集族，使得这个族的每个成员的极限相对频率都存在，但这个族不是一个代数。换言之，$\lim(t(A,\ n)/n)$ 和 $\lim(t(B,\ n)/n)$ 可能存在于 s 中，但 $\lim(t(A\cup B,\ n)/n)$ 不可能存在。下面的例子归功于赫尔曼·鲁宾(Herman Rubin)(口头交流)。

考虑成对地选择元素 00，11，01，10，例如，00，11 和 01，10，在一定程度上，与其极限相对频率不存在的特定序列相似，即，用 2 的乘方的术语来说，

$$s=(00,\ 01,\ 10,\ 11,\ 00,\ 11,\ 00,\ 01,\ 10,\ \cdots)$$

因此，

$$s_n=\begin{cases}00,\ \text{如果 } n \text{ 是奇数，而且使得 } 2^{k-1}\leqslant n<2^k \text{ 的 } k \text{ 是奇数，}\\ 11,\ \text{如果 } n \text{ 是偶数，而且使得 } 2^{k-1}\leqslant n<2^k \text{ 的 } k \text{ 是奇数，}\\ 01,\ \text{如果 } n \text{ 是偶数，而且使得 } 2^{k-1}\leqslant n<2^k \text{ 的 } k \text{ 是偶数，}\\ 10,\ \text{如果 } n \text{ 是奇数，而且使得 } 2^{k-1}\leqslant n<2^k \text{ 的 } k \text{ 是偶数。}\end{cases}$$

我们把 A 定义为某一项的第一个元素为零的事件，B 定义为这一项的第二个元素为 1 的事件。因此，

$$A=\{00,\ 01\}$$

$$B=\{01,\ 11\}$$

$$A\cap B=\{01\}$$

为此，

171

$$P(A)=\frac{1}{2}\quad \text{和}\quad P(B)=\frac{1}{2}$$

至于 n，我们甚至有，对于每个 $n>1$

$$\frac{t[(A\cap B),\ 2^n-1]}{(2^n-1)}=\frac{1}{3}$$

因而,对于 n 是奇数,

$$\text{当}\ n\rightarrow\infty\ \text{时},\frac{t[(A\cap B),\ 2^n-1]}{(2^n-1)}\rightarrow\frac{1}{6}$$

因此, $P(A\cap B)$ 不存在,并且很容易推出 $P(A\cup B)$ 也不存在。

冯·米泽斯。我们的下一个主要问题是描述真正的随机型序列。第一种系统的努力可以说成是来自冯·米泽斯(von Misses)的重要工作(1919)。

冯·米泽斯在他关于随机序列的(集体的)定义中用了两个直观的原理。一个序列 s 是一个随机序列,当且仅当,(i) 在 $R(s)$中每个实数 r 都有一个极限相对频率,以及(ii) 如果一位赌徒使用了根据任意的算术规则(比如,只在奇数的情况下打赌)来确定是否打赌(在给定的 $r\in R(s)$ 内)的一个系统,那么在他选择的序列中 r 的极限相对频率等于在整个序列中 r 的极限相对频率。

第二个原理是著名的排除赌博系统原理,尽管它的名字有些误导,因为它并不关心所有的赌博系统(例如,每次都使你的赌注加倍)。我们将会看到,为这个原理提供一种适当的形式重构是很不容易的。我们从几个预备定义开始。

一个序列是一个*单调增加的正整数序列*,当且仅当,它的值域 $R(s)$是一个正整数的集合,而且,如果 $m<n$,那么 $s_m<s_n$。一个序列 s 是 t 的一个*子序列*,当且仅当,t 是一个序列,并存在着一个 r,使得 r 是一个单调增加的正整数序列,以及 $s=t\circ r$。

定义 3. 一个选址规则(place-selection rule)是一个正整数的单调增加序列。

从直观上看,一位赌徒总是(有意或无意地)根据某个选址规则进行赌博。

定义 4. 设 s 是一个序列,再设 $\mathfrak{J}$ 是在 $R(s)$ 上的一个集合代数。于是,s 是相对于 $\mathfrak{J}$ 的第一种意义上的一个随机序列,当且仅当,

(i) 对于 $\mathfrak{J}$ 中的每个 A,在 s 中 A 的极限相对频率存在,

(ii) 对于 $\mathfrak{J}$ 中的每个 A 和每个选址规则 g,在 $s \circ g$ 中 A 的极限相对频率存在,而且,在 $s \circ g$ 中 A 的极限相对频率的值,等于在 s 中 A 的极限相对频率的值。

不幸的是,这个定义太强。下列定理表明,在定义 4 的意义上,根本没有非平凡的随机序列。

定理 4. 设 s 是一个序列,再设 $\mathfrak{J}$ 是 $R(s)$ 上的一个集合代数。如果 s 是相对于 $\mathfrak{J}$ 的第一种意义上的一个随机序列,那么对于 $\mathfrak{J}$ 中的每个 A, 172

$$P(A) = 0 \quad 或者 \quad P(A) = 1$$

(也就是说,在第一种意义上,根本没有非平凡的随机序列。)

证明: 假定 A 在 $\mathfrak{J}$ 中,而且,$P(A) \neq 0$, $P(A) \neq 1$。于是,我们知道,A 在这个序列中出现无穷多次。因此,对于由出现 A 的那些项所组成的子序列来说,有一个选址规则 g,使得 $s \circ g$ 是那个序列。但是,在 $s \circ g$ 中 $P(A)$是 1,与我们在第一种意义上定义的随机序列相反。

定义 4 在形式上是精确的,但它的很一般的选址规则概念,在任何意义上,都没有捕获到冯·米泽斯的排除赌博系统原理的直观思想,因为在选址规则的定义中根本没有条件要求,(选

择的)第 n 项只是 n 和前一项 $n-1$ 的结果的一个函数。下一个定义试图表达这种思想。

设 I^+ 是正整数集合，设 S 是由 0 与 1 组成的所有序列的集合，即其值域是集合 $\{0, 1\}$ 的所有序列的集合，而且，对于在 S 中的任何一个序列 s 来说，设 $[s]_n$ 是在 S 中它们的前 n 项等于 s 的所有序列的集合。换言之，

$$[s]_n = \{t: t \text{ 在 } s \ \& \ (\forall m \leqslant n)(s_m = t_m) \text{ 之内}\}$$

为了尽可能地捕获到这种直观思想：第 n 次试验用的选址规则应该只依赖试验次数和前一次 $n-1$ 的结果，我们定义

定义 5. 设 ϕ 是把集合 $S \times I^+$ 映射到集合 $\{0, 1\}$ 的一个函数，于是，ϕ 是一个严格的选址规则，当且仅当，只要 $[s]_{n-1} = [t]_{n-1}$，对于在 S 中的任何一个 n 和任何两个序列 s 与 t，

$$\phi(s, n) = \phi(t, n)$$

我们用这个严格的选址规则的定义来理解一个新的随机性的意义。乍一看，定义 5 似乎是对定义 3 的过分宽泛的一般性的一个明确改进。函数 ϕ 的作用是告诉赌徒，何时打赌，恰好像在定义 3 意义上的选址规则的情况一样。在所预期的解释中，只要 ϕ 的值是 1，赌徒就应该下赌注，正如下列定义中表明的那样：

定义 6. 由 0 和 1 组成的一个序列 s 是第二种意义上的一个随机序列，当且仅当，

(i) 在 s 中 1 的极限相对频率是存在的，

(ii) 对于每个严格的选址规则 ϕ 来说，在子序列 $s \circ \phi_s$ 中 1 的极限相对频率是存在的，而且，1 在 $s \circ \phi_s$ 中的极值与在 s 中的极值是一样的。

不幸的是，乍一看，在第二种意义上的随机性的印象是错误的。它不代表对第一种意义上的随机性的任何一种真正的改进，因

为每个单调增加的正整数序列都正好是某一严格的选址规则 ϕ_s。与定理 4 相对应,我们因而有下列定理:

定理 5. 第二种意义上,根本没有非平凡的随机序列。

以任何一种简单的方式来改进第二种意义上的随机性,似乎没有太大的希望。我们的困难是,对子序列或选址规则的显明的集合论构造的无限制地使用,总是以这些规则的完整的经典集合所告终,在定理 4 和定理 5 的观点中,这个经典集合非常之大。定理 4 和定理 5 所造成的相对频率理论的深层次的困难,并不总是受到这种理论的拥护者们的重视,他们一直不情愿在纯字面的意义上接受他们自己的思想。 173

丘奇。我们现在转向解决所遇到的困难的最佳进路。基本思想归功于丘奇(Church,1940)。我们最后把选址规则限定为是递归规则。尽管为了使这种发展成为系统的和形式的,需要几个技术性的定义,但是,我们在这里给出我们对随机性的最终的形式定义,此外,如果有些读者希望跳过下面这个定义的相当技术性的讨论,可以直接转到定理 6。我们用了前面第 3 章的第 5 节中所引入的部分递归函数的概念。

定义 7. 一个由 0 和 1 组成的序列 s 在丘奇的意义上是一个随机序列,当且仅当,

(i) 在 s 中 1 的极限相对频率存在,

(ii) 如果 ϕ 是把正整数集合映射到 $\{0, 1\}$ 的任意部分递归函数,而且,如果 b 是序列

$$b_1 = 1,$$

$$b_{n+1} = 2b_n + s_n,$$

那么在子序列 $s \circ \phi \circ b$ 中 1 的极限相对频率存在,而且,1 在子序列 $s \circ \phi \circ b$ 中的极限相对频率的极值等于 1 在 s 本身中

的极值。

正如下列例子所说明的那样,序列 b 的目的是,把一个变量数可变的函数描述成一个变量的函数,正如已经评论的那样,因为冯·米泽斯原理的意图是:在第 n 种场合下赌注的决定,应该只是前一次 $n-1$ 的结果和数 n 的函数。然而,这意味着,选址函数的自变量的个数是随着每一个 n 的变化而变化的,在把 ϕ 描述为部分递归的时,这在技术上是困难的。因此,我们引入序列 b。我们通常写作 $\phi(b_n)$,而不是写作 $(\phi \circ b)(n)$。与前面一样,在这个定义中函数 ϕ 的作用是告诉赌徒何时下赌注,而不是什么事件上或多少事件上下赌注。当 ϕ 的值等于 1 时,赌徒应该下赌注。下面的例子应该使这一切变得更加清楚了。

定义 7 的 ϕ 函数的第一个例子。 选择 ϕ 是由每个奇数项组成的子序列。从直观上看,我们在第 n 种场合下赌注,当且仅当,$\phi(b_n)=1$。我们希望找到一个 ϕ,使得对于每个 k, $\phi(b_{2k})=0$ 和 $\phi(b_{2k+1})=1$。首先,让我们看一下相对于由 0 和 1 组成的任何一个可能序列的 b_n,我们看到,

174

$$b_1 = 1,$$

$$b_2 = 2b_1 + s_1 = \begin{cases} 2;当\ s_1 = 0\ 时 \\ 3;当\ s_1 = 1\ 时 \end{cases}$$

$$b_3 = 2b_2 + s_2 = \begin{cases} 4;当\ s_1 = 0\ 时, \\ 5;当\ s_1 = 0\ 和\ s_2 = 1\ 时, \\ 6;当\ s_1 = 1\ 和\ s_2 = 0\ 时, \\ 7;当\ s_1 = 1\ 和\ s_2 = 1\ 时 \end{cases}$$

一般情况下,b_n 的可能值恰好是使得 $2^{n-1} \leqslant k < 2^n$ 的整数 k。例如,我们知道,

$$2^{1-1} = 2^0 \leqslant b_1 < 2^1,$$

$$2^{2-1} = 2^1 \leqslant b_2 < 2^2,$$

$$2^{3-1} = 2^2 \leqslant b_3 < 2^3。$$

的直观思想是，不顾实际结果，来构造它们，如果 $n \neq m$，我们总是有 $b_n \neq b_m$。

对只由 0 组成的序列来说，b_n 的值是 2^{n-1}，而对于只由 1 组成的序列来说，b_n 的值是 $2^n - 1$。选择每个奇数项的 ϕ 如下：

$$\varphi(a) = \begin{cases} 0, \text{如果使得 } 2^{k-1} \leqslant a < 2^k \text{ 的惟一的 } k \text{ 是奇数} \\ 1, \text{如果使得 } 2^{k-1} \leqslant a < 2^k \text{ 的惟一的 } k \text{ 是偶数} \end{cases}$$

例如，

$$\phi(b_1) = \phi(1) = 1 \text{ 既然 } k = 1$$

$$\phi(b_2) = \begin{cases} \phi(2) = 0 \text{ 既然 } k = 2 \\ \phi(3) = 0 \text{ 既然 } k = 2 \end{cases}$$

上面的例子只依赖于整数 n。

定义 7 的 ϕ 函数的第二个例子。 这个例子依赖于由 1 和 0 组成的序列的结果。对于相信出现零会使他走运的一位赌徒来说，我们希望找到一个 ϕ，使得 $\phi(b_n) = 1$，当且仅当，$s_{n-1} = 0$。即他在每次出现零之后都下赌注。这里，适当的 ϕ 是，

$$\phi(a) = \begin{cases} 1, \text{如果 } a \text{ 是偶数} \\ 0, \text{如果 } a \text{ 是奇数} \end{cases}$$

定义 7 的 ϕ 函数的第三个例子。 对于只有当前面的两种结果是零时才下赌注的人来说，找到适当 ϕ，

$$\phi(a) = \begin{cases} 1, \text{如果 } a \text{ 可被 4 除尽，} \\ 0, \text{其他情况} \end{cases}$$

一般方法。我们现在能推导出一个一般方法，来为下列赌徒找到适当的ϕ：他只在出现了一个特定的有穷序列$(t_1, t_2, \cdots, t_k)$之后，才下赌注，其中，t_i 的取值为 0 和 1。①

175 更确切地说，我们希望 $\phi(b_n) = 1$，当且仅当，

$$s_{n-k} = t_1,$$
$$s_{n-k+1} = t_2,$$
$$\cdots$$
$$\cdots$$
$$\cdots$$
$$s_{n-1} = t_k。$$

根据 b_n 的定义，

$$b_{n-k+1} = 2b_{n-k} + s_{n-k} = 2b_{n-k} + t_1,$$
$$b_{n-k+2} = 2b_{n-k+1} + s_{n-k+1} = 2(2b_{n-k} + t_1) + t_2,$$
$$= 2^2 b_{n-k} + 2t_1 + t_2,$$
$$b_{n-k+3} = 2b_{n-k+2} + s_{n-k+2} = 2(2^2 b_{n-k} + 2t_1 + t_2) + t_3,$$
$$= 2^3 b_{n-k} + 2^2 t_1 + 2t_2 + t_3,$$
$$\cdots$$
$$\cdots$$
$$\cdots$$
$$b_n = 2b_{n-1} + s_{n-1} = 2^k b_{n-k} + 2^{k-1} t_1 + \cdots + 2t_{k-1} + t_k$$
$$= 2^k b_{n-k} + \sum_{i=1}^{k} 2^{k-i} t_i。$$

但是，根据 b_n 的定义，b_{n-k}一定是一个正整数。因此，我们

---

① 这种推导归功于雅各布(Richard L. Jacob)(口头交流)。

希望的 ϕ 是，

$$\varphi(a)=\begin{cases}1, \text{如果存在着一个正整数 } m, \text{使得 } a = 2^k m + \sum_{i=1}^{k} 2^{k-i} t_i \\ 0, \text{其他情况}\end{cases}$$

事例：为只在出现序列(0, 1, 1)后才下赌注的赌徒寻找 ϕ。这里，$k=3$, $s_{n-3}=t_1=0$, $s_{n-2}=t_2=1$, $s_{n-1}=t_3=1$。

$$\begin{aligned}2^k m + \sum\nolimits_{i=1}^{k} 2^{k-i} t_i &= 2^3 m + 2^2 t_1 + 2^1 t_2 + 2^0 t_3 \\ &= 2^3 m + 2^2 \cdot 0 + 2^1 \cdot 1 + 2^0 \cdot 1 \\ &= 8m + 3\text{。}\end{aligned}$$

因此，我们想要的 ϕ 是

$$\varphi(a)=\begin{cases}1, \text{如果存在一个正整数 } m, \text{使得 } a = 8m+3 \\ 0, \text{其他情况}\end{cases}$$

我们也能以相当不同的方式运用我们推导出的 b_n 的表达式。考虑 $k=n-1$ 的情况。于是，$b_{n-k}=b_1=1$ 和 $t_i=s_{n-k-1+i}=s_i$。因此，

$$b_n = 2^k b_{n-k} + \sum_{i=1}^{k} 2^{k-i} t_i = 2^{n-1} + \sum_{i=1}^{n-1} 2^{n-1-i} s_i$$

现在，如果我们有 b_n 的值，那么我们能只通过把 b_n 表示为 2 的乘方之和，找到 n 的值，也找到序列 s 的前 $n-1$ 项的值。 176

事例：我们有 $b_n = 107 = 64+32+8+2+1 = 2^6+2^5+2^3+2^1+2^0$。因此，$n-1=6$，所以，$n=7$。现在，如果我们把 b_n 写成下列形式，

$$b_7 = 2^6 + 2^5 \cdot 1 + 2^4 \cdot 0 + 2^3 \cdot 1 + 2^2 \cdot 0 + 2^1 \cdot 1 + 2^0 \cdot 1$$

那么我们能立即写出这个序列的前六项，因为他们是 2 的乘方

的系数：

$$s = (1, 0, 1, 0, 1, 1, \cdots)$$

值得关注的是，如果我们放弃 ϕ 是部分递归的必要条件，那么我们能找到一个 ϕ，使得 ϕ 选择一个它的所有项都是 1 的子序列。例如，考虑如下：

$$\phi(a) = s_{\mu}(a)$$

这里，$\mu(a)$等于使得 2^m 大于 a 的最小的正整数 m。细节留给读者。

让我们现在转向丘奇关于 ϕ 是部分递归的建议。他的解决办法取决于 ϕ 是有效地可计算的必要条件。一个函数的有效的可计算性直观上意味着，如果我们已知该函数的自变量，那么我们就有在有限的步骤内计算该函数的一个代数。最后定义的函数 ϕ 显然不是有效地可计算的，因为已知一个正整数 x，根本没有任何办法计算 $\phi(x)$。我们首先必须给出序列 s。

运用从分析和概率理论得出的结果，可以证明下列定理。

定理 6. *存在着在丘奇意义上的非平凡的随机序列。的确，已知在 0 和 1 之间的任意数 a，存在着在丘奇的意义上由 0 和 1 组成的一个随机序列，使得 1 的极限相对频率是 a。*

正如丘奇(1940)所强调的那样，定理 6 的证明依赖于强有力的非构造的证明方法，这是相对频率理论受到的主要形式批评之一；即，概率中的基本组合问题依赖于有力的、非初等的分析方法。例如，在把一枚公平的硬币投掷三次，至少得到两次正面的概率取决于定理 6。通过运用条件断言，能够避免依赖于这个定理，这是真的，但核心事实仍然是，要求像在日常经验中那样普遍存在的概率概念，依赖于甚至连许多数理统计学家都不熟悉的这样一个详细阐述的概念框架，似乎是很不令人满意的。例如，统计学家最熟知的处理这些问题的方法也许是沃尔

德(Wald)的处理方法(1936),从数学的观点来看,这种处理方法是很不令人满意的,因为没有确切地定义,他要求一个选址规则在一个固定的逻辑系统内是可定义,是什么意思。①

从形式的观点来看,历史上对随机序列的这种描述的最重要的批评可能是维尔(Ville)的批评(1939, pp. 58-63)。他表明,冯·米泽斯—丘奇的随机序列概念包括由1和0组成的二值序列,其中,该序列的所有开始的部分包含的1的个数大于等于0的个数。因为任何一个赌场的老板或赌徒会马上明白,在1上下赌注总是能赢的建议之意义上,这样的序列是不公平的。维尔还用鞅(martingale)的重要概念来强化一个公平赌博系统的概念。② 如果一位赌徒预计下一次赌博的运气与他刚结束的赌博的实际运气一样,一个赌博系统就是一个鞅。更技术性地说,在没有完全明确所有假设的条件下(参见 Feller 1966, p. 210 ff),随机变量$\{\boldsymbol{X}_n\}$的一个序列是一个鞅,当且仅当, 177

$$E(\boldsymbol{X}_{n+1} \mid \boldsymbol{X}_n, \boldsymbol{X}_{n-1}, \cdots, \boldsymbol{X}_1) = \boldsymbol{X}_n$$

还有更多的要求是,随机变量$\{\boldsymbol{X}_n\}$的一个序列是**绝对公平的**,当且仅当,

$$E(\boldsymbol{X}_1) = 0$$

而且,

$$E(\boldsymbol{X}_{n+1} \mid \boldsymbol{X}_n, \boldsymbol{X}_{n-1}, \cdots, \boldsymbol{X}_1) = 0$$

(我已经用测量理论的术语,而不是集合的术语,阐述了这些条

---

① 也应该注意到,科普兰(Copeland)的容许数概念(1928)和赖兴巴赫(Reichenbach)的正常序列的等价概念(1932)预料到了丘奇的分析,但是由于这里不考察的技术性的理由,不是令人满意的。

② 鞅这个术语是由维尔提出的,但是,在莱维(P. Levy)和有可能其他人的早期著作中已预料到这个概念。

件,因为这样的测量理论的语言现在是概率的通用语言,与对待概率基础的态度无关。)

至于反对维尔的批评,为冯·米泽斯的极力辩护,可参见范·兰巴根的文章(van Lambalgen,1987a, pp. 43－54; 1996)。

更加新近的工作从一个不同的方向巩固了随机序列的描述。最重要的观点是序列的可计算复杂度作为一种不同的概念方式判断其随机性的观点。详细的历史太复杂,这里无法展开。但要点如下。1964年,柯尔莫哥洛夫(1965)引入了符号的一个有穷序列的复杂度概念,这些符号来自一个固定的有穷字母表,例如,1和0。他把这样一个序列的复杂度定义为,对这个序列进行编码所要求的最短程序的长度,正如下一节更详细地讨论的那样,最短是相对于给定的计算机和程序语言而言的。柯尔莫哥洛夫思想的自然设置是有穷序列,而不是无穷序列。下一节还要涉及马丁勒夫(Martin-Löf)的相关工作。

的确,冯·米泽斯的相对频率理论最终似乎是两头失误。一方面,概率的数学理论与本章第1节略述的标准的测量理论的进路相比,是很不方便的和不自然的。还有,尽管它在技术上很复杂,但也没有为研究有穷序列提供一个适当的应用理论。关于概率基础的偶然的哲学著作一直没有充分地强调,在为适合研究复杂概率现象的这种相对频率理论提供一个完备的和字面的陈述时,所包含的数学困难。当然,就我所知,还没有人从相对频率的立场认真地分析过随机过程。一旦像在随机过程中那样,承认一个事件序列中的依赖性,那么除了独立事件的特殊情况之外,假如冯·米泽斯的最初思想得到了实现,随机性的适当定义在其复杂度方面就变得更加令人恐惧。困扰随机过程理论的数学问题,在一个更简单的测量理论的框架内,过分充分地证明,保留在这个框架内并避免相对频率理论的不必要的复杂

178

性是有道理的。相对于应用问题来说,明显的事实是,应用概率理论研究有穷序列,而不是研究无穷序列。应用所需要的是在有穷序列中更精确的和更满意的随机理论。

最后,在说明为什么冯·米泽斯—丘奇的相对频率理论现在没有被普遍地公认为是概率的理论基础的问题时,我们应该重新强调上面所阐述的丘奇的重要观点。似乎明显的是,研究作为定义 7 所定义的随机序列所需要的微妙的非构造方法,肯定似乎不能满意地解决掷骰子或掷硬币的简单的直观问题,这些问题在经典理论中得到了这样一种自然的处理。打算替代经典理论的任何一个理论,为了最终被判定为是令人满意的,必定提供一种比较简单的方法处理赌博游戏。概率的相对频率理论或至少是当前的概率观点,不可能以完全形式的方式做到这一点。范·兰巴根(1987a)强调了作为物理学上最有吸引力的替代选择的倾向性解释的重要性(参见,本章的第 6 节)。

同样重要的是,当把柯尔莫哥洛夫等人提出的有穷序列的复杂度概念扩展到无穷序列时,那么这种随机性的定义要比冯·米泽斯—丘奇的定义更严格,所以,有些序列在冯·米泽斯意义上是随机的,而在柯尔莫哥洛夫意义上却并非如此。

相关的大量而复杂的文献资料,强调了这里给出的这种分析的有限特征。至于一种极好的综述,读者可参考范·兰巴根(1987a, b)的文章;不幸的是,他的一篇文章(1987a)没有得到应有的理解,但也可参见他的著作(1990,1992 和 1996)。关于更一般的参考文献,可参见下一节最后提供的那些文献资料。

5.4 随机有穷序列

从无穷序列转到有穷序列的一个主要论据是,在科学实践

中,只遇到有穷序列。柯尔莫哥洛夫的复杂度定义(1963,1965)的一个中心思想是,一个有穷序列,如果它的描述是关于相同长度的,那么它就是随机的。还有一个不同的直观思想是,给出随机性的一个程序描述。一个有穷序列,如果它是由一个随机过程产生的,它就是随机的。在一定程度上,实际的有穷序列和程序之间的对比反映了集合进路与测量理论进路之间的对比,就像在关于这些问题的著名争论中所反映的那样(Doob 1941; 179 von Mises 1941)。[①] 一个典型的数学措施是用函数表示程序,而且,现有的一个文献是关于随机函数的(例如,Goldreich, Goldwasser and Micali 1986)。这些作者的直观思想是,一个函数,如果寻找该函数值的时间多项式的算法,对于根据这种算法选择的自变量来说,根本无法把对该函数是有效的一种计算与只表示独立地掷硬币的结果的一种计算区别开来,这个函数就是**多重随机的**。

尽管细节并不清楚,但是,从这个定义似乎得出,塔斯基的初等代数的决策程序是多重随机的,因为它是一种时间指数算法。显然,多重随机性的思想绝不是一个绝对的随机性概念,因为我们明确地认为不是直观上随机的过程,也满足这个定义。多重随机函数定义的另一个不能令人满意的特征是,它把随机性概念简化为取决于本身在直观上是一个随机过程的独立地掷硬币的结果。从更广泛的哲学观点来看,这样的一种简化不是很有帮助,因为我们的目标之一,是描述何时把一个掷硬币的过程或一个实际的投掷序列看成是随机的。

随机函数的另一个直观思想是支持随机变量的标准概念的

---

① **测量理论的**这个术语恰好指由柯尔莫哥洛夫(1933)提出的和在第1节的定义2中运用的形式概率测度的类型。

思想。我们预计在一个函数的值中有无法说明的可变性。从这里的思想如何接近于能解决某些问题的标准的统计实践来看，这是显然的。马丁勒夫对随机(无穷的)序列的描述完全是沿着这些思路进行的，而且，是对上一节观点的部分发展。马丁勒夫(1966)的核心思想是，一个序列，如果通过任何有效的统计检验，在任意小的意义层次上，都不可能被拒绝，那么，相对于一种测量而言，它就是随机的。从对无穷序列的这种一般描述转到对有穷样本的显著性检验，是很有道理的。尽管有问题需要克服，但似乎对我来说，如果非绝对的结果是可接受的，那么，沿着有穷样本的显著性检验的思路，提出有穷随机序列的理论存在的基本障碍，不比采取柯尔莫哥洛夫的路线存在的基本障碍更多。注意，在这两种思路之间，实际上有着概念上的不同。后一种进路，即经过显著性检验的思路，我们可能称为随机进路，而柯尔莫哥洛夫采取的路线，我们依据文献，称为复杂度进路。尽管我们指望这两条进路能渐进地接近于某些共同的结果，但是，根本没有先天的理由指望两者得出相同的结果，除非在抓住我们似乎具有的随机性的直觉时，一条进路显然是有缺陷的。关于这方面的更多细节，参阅下面的定理 4。

柯尔莫哥洛夫复杂度。柯尔莫哥洛夫最早的陈述之一是：有必要把相对频率作为概率的一个关键问题来解释，这种陈述是在建议对有穷序列的随机性进行统计检验的语境中作出的。他的所述如下。

> 我已经表达了这种观点(参见，Kolmogorov 1933，第一章)，把概率的数学理论结果应用于真正的“随机现象”的基础，一定依赖于**概率的频率概念**的某种形式，这个概念的不可避免的本性是由冯·米泽斯以一种令人兴奋的方式确立的。然而，长期以来，我一直持有下列观点： 180

> 1. 以当试验次数增加到无穷时的极限频率概念为基础的频率概念,对证明把概率理论的结果应用于我们总是处理有穷次试验的真正的实践问题,没有任何贡献。
>
> 2. 应用于大量但有穷次试验的频率概念,在纯数学的框架内,不承认一种严格的形式阐述。
>
> 因此,我有时提出了这样的频率概念:它包括,在没有"足够大的"系列的精确定义……的条件下,自觉运用"实践的可靠性"、"在一个长系列的试验中频率的近似稳定性"这些确定的但不严格的形式思想。
>
> ……我还坚持上面提到的两种观点中第一种观点。然而,至于第二种观点,我已经意识到,在一个大的有穷人口中一种特性的随机分布概念可能有一个严格形式的数学阐明。事实上,我们能够表明,在足够大量的人口中,这种特性分布可能是这样的:**当选择子人口的规律足够简单时**,这个特性出现的频率,对于所有足够大的子人口来说,几乎是一样的。当全面发展这样一个概念时,需要引入对运算规则的复杂度的一种测量。我在另一篇文章中建议讨论这个问题。然而,在现在的这篇文章中,我将运用的事实是,不可能有**大量的简单的运算规则**。

柯尔莫哥洛夫从来没有放弃他的信念:有穷序列,而不是无穷序列,为随机性分析提供了适当的科学背景。

我现在转向柯尔莫哥洛夫的形式的复杂度理论(1963,1965)。(索罗门诺夫(Solomonoff)独立地提出了相同的概念,蔡廷(Chaitin)的工作(1966,1969)也与此相关。)

我们大家都承认,有些序列是很简单的。例如,现在考虑长度为 n 的序列,只由 1 组成的一个序列或只由 0 组成的一个序

列，显然在结构上是非常简单的。我们对简单的周期序列也有同感，例如，由简单地交替 1 和 0 组成的序列。问题是，如何以一般的方式定义有穷序列的复杂度。柯尔莫哥洛夫的卓越的直觉是，通过产生一个序列所需要的最小程序的长度的方式，来测量这个序列的复杂度。在考虑描述柯尔莫哥洛夫的复杂度所需要的形式定义之前，提出这些定义为什么本来就是复杂的警告，是有用的。这种直觉想法是相当自然的。我们似乎能挑选出一些简单的通用图灵机，或者一些简单的寄存器机，用作我们有效地计算任何一个有穷序列的复杂度的固定计算机。然而，像阅读编码各种复杂指令所用的程序技巧表明的那样，问题要复杂许多。因为我们坚持长度最短的程序，我们将会面对任何能力的任何计算机的一个常见问题。处理承认个别序列的独特特征的程序并不简单。所有这些要点的结果是，柯尔莫哥洛夫的复杂度概念不是一个有效地可计算的函数，即它不是一个部分递归函数。如果我们思考集合的核心特征，那么它不是有效地可计算的，也许不会令人感到惊奇，尽管朴素的直觉可能建议它应该是可计算的。 181

由于想到这些复杂度，我们转向形式描述。首先，我们将把通用计算机的合适的一般类，比如，第 3 章描述的无限寄存器机或通用图灵机，看成有能力计算任何部分递归函数。

设 x 是一个 1 和 0 的一个有穷的二元序列，设 U 是一个通用计算机，设 $U(p)$ 是 U 在运行程序 p 时 U 的输出，再设 $l(p)$ 是程序 p 的长度。

定义 1. **相对于 U 的一个字符串 x 的柯尔莫哥洛夫复杂度被定义为，作为输出和停止打印 x 的所有程序所接受的长度最短的程序，即用符号表示为**

$$K_U(x) = \min_{\{p:\, U(p)=x\}} l(p)。$$

这种测量复杂度的最基本特性是，任何一个其他的通用计算机 V 都将提供一个对 x 的复杂度的测量，x 的复杂度在独立于 x 的常数范围内，在这种意义上，它是通用的。

定理 1.（普遍性）。设 V 是任何一个别的通用计算机。于是，存在着一个常数 $c(V)$，使得对于任何一个有穷的二元序列 x

$$K_U(x) \leqslant K_V(x) + c(V)。$$

证明的思路是，既然 U 是一个通用计算机，它就能根据一个长度 $c(V)$ 的一个有穷程序模拟 V，因而总是满足这个不等式。

在这个定理的观点中，我随后不再提到 U，由于模拟程序的长度，即使是短程序，也可能有很大的不同。长的复杂序列的测量，不应该是这种情况。

显然，上述定义 1 实际上不是我们想要的，大体上说，复杂度显然随着序列长度的增加而增加。我们希望能够作出的判断是，对于固定长度的序列来说，它们的相对复杂度，而且，把足够复杂的那些序列判断为是随机的。这意味着，在给出关于序列的某些信息的地方，我们实际上需要一个复杂度的条件概念。

因此，定义 1 的一个重要修改是，使符号串 x 的已知长度 $l(x)$ 的复杂度的测量条件化。用符号表示为

$$K_U(x \mid l(x)) = \min_{\{p:\, U(p,\, l(x)) = x\}} l(p)。$$

这里是与一个序列的复杂度及其长度相关的一个自然结果。

定理 2. 一个有穷的二元序列的条件复杂度小于它的长度加独立于 x 的一个固定常数。

证明取决于拥有一个简单的打印指令的等价性：

$$打印\ x_1, x_2, \cdots, x_{l(x)}。$$

定理 3. 柯尔莫哥洛夫的复杂度测量 K 不是一个可计算的函数。

证明: 证明依赖于这样的事实：当我们在寻找最短的程序时考察程序 P,我们不能有效地确定,一个已知程序 P 是否是一个候选者,因为我们无法以一种有效的方式确定,一个已知程序 P 是否停止。如果对于输入 s 来说,它确实停止了,那么它就是一个候选者;如果它没有停止,它就不是一个候选者。记住:是有效地可计算的,我们必须能够提供一个运算规则,使得对于任何一个程序和任何一个序列 s 来说,我们能计算这个程序是否会停止。根据停止问题的著名的不可判定性,我无法做到这一点。显然,对于一个特殊序列来说,我们能够作出特殊的论证,但这并不意味着复杂度测量不是可计算的。我们打算提供一个适用于所有情况的运算规则。 182

通用概率。设 p 是一个程序,使得 $U(p)=x$。如果 p 是通过掷一枚公平硬币 $l(p)$次的一个伯努利过程随机地产生的,那么 p 概率是,

$$2^{-l(p)}$$

因此,对于 x 的通用概率 $P_U(x)$来说,我们是求随机产生的概率之和,即,通过掷硬币,至少程序 p 之一,使得 $U(p)=x$:

$$P_U(x)=\sum_{\{p:\,U(p)=x\}}2^{-l(p)}$$

尽管我没有尽可能形式地给出这个定义,但是,应该明确的是,通用概率 $P_U(x)$是决不依赖于有穷序列 x 的柯尔莫哥洛夫复杂度的一个概率的测量理论概念。

这是与两个概念相关的下列定理的意义,即一个有穷序列 x 的通用概率和它的柯尔莫哥洛夫复杂度。

定理 4. 存在着一个常数 c,使得对于每一个有穷的二元序

列 x,

$$2^{-K(x)} \leqslant P_U(x) \leqslant c2^{-K(x)}$$

完全证明这个定理太复杂,这里无法提供。事实上,这一节的整个进展,对于全面掌握这种思想来说,没有所希望的那么详细。幸运的是,在文献中有许多好的详细阐述:科弗(Cover)和托马斯(Thomas)(1991);科弗、高奇(Gacs)和加里(1989);李和维塔尼(Vitanyi)(1993);马丁勒夫(1969)以及兹沃金(Zvonkin)和莱文(Levin)(1970)。

这一节最终仍然没有解决的问题是,我们应该如何定义有穷的二元序列的随机性概念。柯尔莫哥洛夫(1968)提出把定理2作为描述一个有穷序列的随机性的一个基础,即一个有穷序列 x 是随机的,如果它的条件复杂度接近于上极限,也就是说,接近于 x 的长度加上独立于 x 的一个常数。柯尔莫哥洛夫指出,这个定义是明显的。它本质上必定是相对的。同样相对的评论适用于对一个有穷序列的随机性的统计检验。正如柯尔莫哥洛夫(1963)评论的那样,必须考虑到检验的次数及其简单性,但还是没有一个绝对的定义。不管怎样,太多的检验和运算上太复杂的检验将拒绝把一个有穷序列看成是随机的。

也许不令人感到惊奇的是,通常承认,只能为无穷序列提供一个绝对的随机性概念,但在实验设计中,或者在相对于一个假设、模型或理论(视情形而定)的实验结果的统计检验中,如果没有直接的科学应用的可能性,那么这样一种理想的结果,在很大程度上是科学理论化的典型。

183

作为概率估算的相对频率。在具体实践中,有经验的统计学家和科学家很好地应用着关于有穷序列的随机性的实用规则。这里是我所说的这种实用主义的一个简单例子。假设在一

个学习实验中,我需要使奖赏 R_1 或 R_2 反应的 200 个试验随机化。奖赏程序与反应无关,它是在任何一个试验中奖赏 R_1 的概率为 1/2 的伯努利程序, 因而也以 1/2 的概率奖赏 R_2(但不能同时奖赏两者)。于是,奖赏序列的总数 2^{200} 是巨大的。为了达到计算的目标,我们希望保留 $p(R_1)=p(R_2)=\frac{1}{2}$ 的伯努利测量,但不明显地影响计算我们能够排除用 2^{20} 等最低复杂度序列。现在,我们通常实际上不计算这种低复杂度序列的集合,而只是拒绝似乎太规则的序列。我讲的这个故事太简单,但其中隐含了一种明显慎重的实用性观点。在一个给定的科学语境中,可能需要或要求提供更仔细的阐述,但在实践中,如果没有对一个给定的有穷序列作出特别复杂的计算,那么通常依靠获得一个低复杂度序列的低概率。

在第 8 节中,来自物理学家的大量引文所强调的更基本的观点是,统计学家和科学家,即使关于概率本性的观点大不相同,也都用从观察的有穷序列获得的相对频率来估算概率、随机变量的矩、理论模型的参量等。事实上,关于概率本性以及概率在科学中所起的作用,还没有一个真正的思想流派只拒绝所观察到的现象的相对频率所起的任何一种重要作用。通常也不能说,概率恰好是从观察得到的相对频率和一定是有穷的数据。

尽管如此,关于相对频率的作用存在着深入的和认真的争论,这种争论至少可追溯到 19 世纪末。自从 20 世纪中叶以来,这种争论成为现代统计学的核心。现在,在本书中分析所有细节,实在太复杂和太精细,但我希望有些段落将表明,争论的一方,即频率论者,正如其名称所建议的那样,如何支持概率的频率表征或解释。[对于一般的科学读者来说,近来对这种争论的一种出色的更详细的说明,可参见鲁阿内(Rouanet)等人的工作(1998)。]

19 世纪的经典争论是下列两种概率之间的争论：一种是表达信念的与认识相关的概率，特别表达关于像明天帕洛阿尔托(Palo Alto)会下雨之类的单个事件；另一种是频率论者的概率，即意指根据观察到的事件的有穷序列估算的概率。这样，**与认识相关的概率**代表了一种主观的观点，第 7 节分析了这种观点，而**频率论者的概率**代表了概率的一种客观的观点。现在，贝叶斯和拉普拉斯被看成是与认识相关的观点的早期支持者，尽管其他人也起到了重要的作用。英国学派的统计学家卡尔·皮尔逊(Karl Pearson, 1857 - 1936)、菲舍尔(R. A. Fisher, 1890 - 1962)以及后来的耶日·奈曼(Jerzy Neyman)和卡尔·皮尔逊的儿子艾贡·皮尔逊(Egon Pearson)促进了在 19 世纪末和 20 世纪初提出的频率论者的观点。除了任何技术细节之外，频率论者的观点拒绝承认任何主观概率在科学中有任何系统的地位。对受到像柯尔莫哥洛夫之类的有影响人物赞同的这种客观观点的更广泛的接受，占有支配地位，而且，到 1950 年左右，奈曼—皮尔逊的统计进路通常被贴上了正统观点的标签。[1] 贝叶斯的观点在 20 世纪下半叶的复兴主要是一个统计故事，但不出乎意料，它是复兴概率的主观理论的来源，第 7 节考察这个问题。

184

5.5 概率的逻辑理论

我分类为主观的经典理论，用拉普拉斯的用语来说的，实际

---

① 关于奈曼—皮尔逊统计进路的令人鼓舞的哲学辩护，参见梅奥(Mayo, 1996)，关于出色的回顾性说明，参见奈曼(1971)。关于这些问题能作出的更多阐述，它们关注的是统计的基础问题，而不是概率的基础问题，因此，就统计学来说，我的评论和参考文献大多数很肤浅。这是令人遗憾的，但一种恰当的处理需要占用太多的空间。

上是建立在我们不知道真正原因的基础上,与这种经典理论或客观的相对频率理论相反,还存在着作为逻辑延伸的现在不是如此杰出的一种重要的概率观。概率的逻辑关系是命题(可算作是证据)和结论(与证据相关的命题)之间的关系。这种逻辑关系,像推理规则一样,是运用标准术语而不是心理学术语进行思考。这种观点的拥护者在19世纪就能找到,但在20世纪[①]这种立场首先由约翰·梅纳德·凯恩斯(John Maynard Keynes)以最持久和最有说服力的方式提出(1921)。

凯恩斯。他确信莱布尼茨(Leibniz)最先相信概率是逻辑学的一个分支这种思想,但我的目标不是追溯凯恩斯之前的历史。[②]正如在第1节已经提到的那样,凯恩斯强烈地提倡,根据命题,而不是事件,来思考概率。在拥护概率的一种逻辑观点的那些人当中,在这个论点上的惟一变化是,选择作为句法对象的语句,还是选择作为语义对象的命题,作为主要的论题。

这里不是引用凯恩斯公理的时候。与他关于概念的许多有趣而有争议的评论相反,这些公理一般说来是无懈可击的。它们主要以有条件的形式来阐述,通常情况是,不同概率理论的形式方面的描述,代表了基本上所有的思想流派都认同的概率的特性。从本章的宗旨来看,不可能为凯恩斯的观点证明一个有趣的表征定理。然而,他的公理有一点是奇特的和值得关注的。他在第一个公理中指出:

> 假定 a 和 h 是命题,或者,命题的合取或命题的析取,

---

① 原文是"in this century",为了明确起见,这里根据语境译成"20世纪"——译者

② 关于莱布尼茨在17世纪对概率思想的贡献的一种信息量大的分析,参见哈金(Hacking, 1975,第十章)。年轻的莱布尼茨是在合法推理的语境中提出他的概率思想的,他自己早期所受的教育是法律,不是数学或哲学。

185 而且，h 是无矛盾的合取，那么在作为结论的 a 和作为前提的 h 之间，存在一种且只有一种概率关系 P。这样，任何一个结论 a，相对于任何一个一致性的前提 h 而言，都具有一种和只有一种概率关系。

(Keynes, 1921: p.135)

正如平常陈述的那样，我们认为，证明存在着这样一种概率关系，而且，可能更强的结果是，这种概率关系是惟一的，这是一个问题。凯恩斯只是假定了这一点，因而，排除对需要什么条件才能确保存在性和惟一性的任何一种令人感兴趣的分析。

凯恩斯的观点的一个令人惊奇的方面是，他坚持认为，不是所有的概率都能够用数值来表示：①

承认不是所有的概率都能用数值来表示这个事实，限制了无差别原理的范围。人们总是一致认为，一种数值测量实际上能在下列那些情况下获得：即还原为一个惟一完备的等概率选择的集合是切实可行的……但是，在数值测量从理论上和实践中都是不可能的地方，承认这个相同的事实，更有必要讨论这样的原理：它们证明在概率之间进行的大概比较是正当的。

(Keynes, 1921: p.65)

杰弗里斯。比凯恩斯更伟大的一位重要历史人物是哈罗德·杰弗里斯(Harold Jeffreys)，②他常常被引证为是一位主观

---

① 凯恩斯的这种态度影响了后来库普曼(B. O. Koopman)关于部分地要求定性概率的更详细的工作(1940 a, b)，尽管库普曼的观点更多地与主观概率的看法相一致，而不是与概率的逻辑理论的看法相一致。

② 杰弗里斯的《概率理论》有四个版本：1939、1948、1961 和 1983。只有 1948 年的第二版是被认真地修改过的。我的引用和引证取自这个版本。

的贝叶斯主义者,即使他显然相信客观概率。[1]他经常被贝叶斯主义者所引用,是因为他在地球物理学中的许多重要应用运用了先验概率。他对这些应用的深奥而巧妙的思考是把他誉为历史上的重要人物的标志。[2]他对客观先验概率的辩护进展不是很顺利。我提到某些特征。他的概率理论似乎最好描述为合理信念的标准理论,因而是演绎逻辑的自然的归纳延伸。杰弗里斯以一种明确的方式指出了这一点。

> 我们必须从一开始就注意到,归纳比演绎更普遍。后者给出的回答限于简单地说"是"、"不是",或者,"它是不可遵守的"。归纳逻辑必须把最后一种选择,即对演绎逻辑不感兴趣的选择,分解为许多其他选择,然后,根据可供利用的证据说出,它们中的哪一种选择是最有理由相信的。完备的证明和反驳只是两种极端情况。任何一种归纳推理在本性上都包括这种可能性:最有可能做出的选择事实上可能是错误的。总是有例外的可能,而且,如果一个理论不能防止有例外,那么,当它不可能是演绎的时候,它将被断言为是演绎的。由于这种额外的一般性,归纳必须包含在演绎中不包括的公设。我们的问题是陈述这些公设。重要的是注意到,这些公设不可能通过演绎逻辑来证明。如果并非如此,归纳就被还原为演绎,这是不可能的。同样,它们不是经验概括;因为它们需要通过归纳得出,而且,这种论证将是循环论证。我们事实上必须把理论的一般规则和经

186

① 关于杰弗里斯的这一小节得益于读加拉沃蒂(Galavotti)的有关杰弗里斯的最近一篇文章(正在印刷中)。

② 杰弗里斯(1948,第二版)针对凯恩斯的论述是这么说的。"这本书充满了令人感兴趣的历史数据,包含了许多重要的批评性评论。它在构造问题上很不成功,因为不愿意推广这些公理,阻止了凯恩斯获得许多重要结果。"

> 验内容区分开来。一般规则是先天命题，被公认为是独立于经验的，而且，通过它们自身不作出任何关于经验的陈述。归纳是把这些规则应用于观察数据。
>
> (Jeffreys, 1948: pp. 7－8)

在紧接着的下面一段陈述了逻辑与数学的紧密性：

> 于是，一般规则的检验，不是任何一种类型的证明。这不是提出异议，因为演绎逻辑的原始命题也不可能得到证明。所能做到的是，尽可能似真地陈述一组假设，然后看一下，它们会把我们带向那里。演绎逻辑和数学基础的最全面发展是《自然哲学的数学原理》，这本书以被接受为是公理的许多原始命题为出发点；如果这些结论是可接受的，那是因为我们愿意接受这些公理，不是因为后者得到了证明。同样的情形也适用于（过去常常适用于）欧几里得几何。当这是纯数学本身的情境时，我们不一定希望证明我们的原始命题。但逻辑学家和数学家的程序提出的许多建议是，我们在陈述原始命题时，有规则引导我们。
>
> (Jeffreys, 1948: p. 8)

杰弗里斯对于判断信念的合理程度的态度，非常像判断任何一个测量程序结果的态度。观察者可以不同，误差可以产生，但正如在物理测量的情形中那样，这并不意味着没有客观的合理性标准。下面是他针对这种论点所指出的。

> 在不同意这种理论结果的个人评价之间的差异，将是一个心理学的论题。在不降低所参考的惟一标准的重要性的情况下，也能承认存在这些差异。据说，这种概率理论，只要得到实验证据的支持，就能被接受；心理学应该发明衡量信念的实际程度的方法，然后，把信念度与理论相比较。

我应该答复说，如果没有分析观察和从分析观察中进行推理的客观方法，我们就不能够解释这些观察中的任何一种观察……没有人说，错误的回答使算术无效，相应地，我们也不需要说，与概率理论不一致的某些推理的事实使该理论无效。足够明确的是，该理论确实表征了日常思想的主要特征。一个形式理论的优势是使人们更容易看到，在任何一种特殊情况下，这些日常规则是否被遵循。

这种区别表明，在理论上，一个概率应该总是能被完全计算出来的。我们再一次从纯数学上有了一种例示说明。在 e 的展开中第 1 000 个数字是什么呢？没有人知道；但这不是说，它是 5 的概率是 0.1。通过遵循纯数学的规则，我们一定能确定概率的大小，而且，这种陈述要么是根据规则推出来的，要么与规则相矛盾；用概率的语言来说，基于纯数学的数据，它要么是一种必然性，要么是一种不可能性。同样，猜测不是概率。概率理论比演绎逻辑更复杂，甚至在纯数学中，我们也必须经常对近似感到满意。 187

(Jeffreys, 1948: pp. 37 - 38)

杰弗里斯的论述是认真地进行案例研究的一个极好来源，而且，对其他概率观或特殊的统计方法的许多批评性评论，也是出色的。下面是批评拉普拉斯的概率定义以及相对频率定义的一大段。

如前所述，当前的大多数统计理论，被认为似乎依赖于这种或那种不同的概率定义，这些定义断言，要避免合理的信念度的概念。它们的目的是减少公设的个数，一个非常值得赞美的目标；如果能够避免这个概念，我们的第一个公理就是不必要的。我的论点是，这个公理是必要的，并且，

在实践中，还没有一位统计学家用到频率定义，而是所有的人都使用了合理的信念度的概念，甚至通常没有注意到，他们正在使用这个概念，而且，由于使用这个概念，他们正在与一开始提出的原理产生冲突。我不是作为对他们结果的一种批评提出这一点的。当他们开始进行具体应用时，他们的实践大多数是非常好的；错误出在规则中。

已经接受的三种定义是：

1. 如果有 n 种可能的选择，p 是真的选择有 m 种，那么 p 的概率被定义为 m/n。

2. 如果一个事件出现的次数非常多，那么当试验次数趋于无穷时，p 的概率是，p 是真的次数与整个试验总数之比的极限。

3. 实际上，假设可能的试验次数是无穷的，那么 p 的概率被定义为，p 是真的次数与总次数之比。

第一个定义有时被称为"经典"定义……第二个定义是维恩极限，它的现代的主要代表人是米泽斯(R. Mises)。第三个定义是"假定无穷总体"，并通常与菲舍尔的名字联系在一起，尽管它在较早的威拉德·吉布斯(Willard Gibbs)的统计力学著作中出现过，吉布斯的"系综"仍然起着幽灵般的作用。这三种定义有时被假设为是等价的，但在数学的意义上，这肯定是不正确的。

第一个定义出现在棣美弗的书的开头。它通常为概率提供一个确定的值；困难是，这个值通常是其使用者马上会拒绝的值。比如，假设我们考虑两个盒子，一个盒子里装有一个白球和一个黑球，另一个盒子里装有一个白球和两个黑球。随机地选择一个盒子，然后，从这个盒子里随机地取出一个球。取出的这个球是白球的概率有多大呢？共有五

个球，其中两个球是白的。因此，根据这个定义，概率是2/5。但大多数统计学的作者，我认为还包括自称是接受这个定义的那些大多数人，将给出 $\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{3}=\frac{5}{12}$。根据当前的理论立刻得出的结论是，这些术语代表了把乘法规则两次应用于提供两个白球中取出每一个白球的概率。这些因此是通过加法规则附加的。但命题不可能被表达为从 12 种选择中析取出 5 种选择……

关于第二个定义和第三个定义，我们必须记得我们关于一个理论的一般标准。这个标准实际上减少了公设的个数了吗？能够把它应用于实践吗？现在，这些定义明显地不满足第二个标准。没有一个概率在实践中已经是或将来永远是通过下列方法估算的：即计数无穷多次的试验，或者，在一个无穷系列中找到一个比率的极限。第一个定义要么提供了一个无法接受的估算，要么提供许多不同的估算，与此不同，这两个定义根本什么也没有提供。只有通过作出这些结果是什么的假设，才能根据这两个定义得到一个确定的值。甚至这种存在性的证明是不可能的。根据极限的定义，如果没有限制事件出现的可能顺序的某一规则，就根本没有极限存在。极限的存在性被米泽斯看成是一个公设，而维恩几乎没有把它考虑为是一个需要的公设。 188

(Jeffreys, 1948: pp. 341 - 342,345)

杰弗里斯的批评是卓越的，但违背了古老的格言：关在玻璃房子里的人不应该掷石头，在这种意义上，他自己对概率的解释相当缺乏明晰性和严密性。在他的论述中(1948 年第二版)和他的小册子《科学推理》一书中，他提供了一些明确的定性的

概率公理，但是，没有一组公理足以证明一个表征定理，他似乎也不承认需要这样的一个定理。

在更一般的表征意义上，正如已经表明的那样，他说出了许多有趣的东西。例如，他提出，从基本意义上看，两种先验概率的分布代表了既熟悉也被别人较早地使用过的总的无知，但他对它们的论证是值得关注的。①

> 我们的第一个问题是，找到一种方式说，当没有一个可能的值需要特别注意时，一个参量的大小是未知的。两种规则好像覆盖了这些最普通的情形。如果这个参量可以在一个有穷的值域中或从$-\infty$到$+\infty$拥有任何一个值，它的先验概率应该被看成是均匀分布的。如果它以这样一种方式产生，可以想象，它可能有从0到∞的任何一个值，它的对数的先验概率应该被看成是均匀分布的。在根据几组不同的参量都同样好地表示一个定律的地方，存在有估算的情形，而且，令人向往的是，有一个规则使得无论选择哪组参量，都会导致相同的结果。否则，我们将再一次陷入任意地运用适合我们爱好的不同规则的危险。现在知道，存在着具有这种不变性的一个规则，它能应用于很广泛的范围，尽管不是应用于全域。
>
> 这些规则的基本作用是，在允许的值域中，提供一种形式的方式来表示对这个参量值的无知。他们没有作出下列陈述：那个参量或相似的其他参量在不同的值域中发生的频率是多少。它们的作用只是尽可能客观地提供使理论开始的形式规则。从先验概率的任何一种分布开始，并通过

① 杰弗里斯的这些先验分布是不恰当先验的著名事例，因为它们加起来的和是∞，不是1。从杰弗里斯下面提供的例子中能够看出，如何在实践中使用它们。

> 逆概率原理对一系列大量数据的说明,我们无论如何能在任何一个指定的知识阶段提出对相应概率的一种解释。关于一直还没有考虑到的中间步骤,根本没有逻辑问题。但在开始时有一个逻辑问题:除了刚才表明的很模糊的知识之外,当我们对这个参量的值一无所知时,我们如何能为先验概率赋值呢?当人们承认,概率只是与合理的信念度相联系的一个数目,并且除了提供一种形式的表示之外,没有其他目的时,答案实际上是足够明确的。如果我们没有与一个参量的实际值相关的任何信息,那么必须选择概率,以表示我们一无所知这个事实。关于这个参量的值,除了根据其本性有可能被限于某些确定的极限内这个空洞的事实之外,一定说不出什么。
>
> (Jeffreys, 1948: pp. 101 - 102)

189

运用这些思想,这里是杰弗里斯对困扰相对—频率者的问题的解决方案。当只给出一种观察时,我们如何能估算概率呢?除了一种观察之外,根据完全无知的假设,人们一定会强调说,杰弗里斯完全理解,他的估算是合理的,但不一定有很高的精确度。

> 几年前,纽曼(M. H. A. Newman)教授向我提出下列问题。一个在国外旅行的人不得不在一个中转站换乘火车,然后,到达他只听说是存在的一个城镇。他对这个城镇的规模没有概念。他看到的第一样东西是,电车的编号是100。他能从这个城镇里的电车的编号推出什么呢?为了达到这个目的,他可能会假定,电车的编号是从1开始连续地增加的。
>
> 这个问题的新颖之处在于,所估计的量是一个正整数,

它的可能值没有明显的上极限。一个均匀的先验概率因此是不可能的。对于没有上极限的一个连续的量来说，dv/v 规则是惟一地满足的，……

(Jeffreys, 1948: pp. 213 - 214)

这里是杰弗里斯的分析，我为了与本章相符合，对其中的符号做了改动。设 $\boldsymbol{X}$ 是一个数值观察的随机变量，设 $\boldsymbol{Y}$ 是这个城镇里电车数的随机变量。于是，$\boldsymbol{Y}$ 的离散值的先验概率是

$$P(\boldsymbol{Y}=n) \propto \frac{1}{n}$$

即与$\frac{1}{n}$成正比。假设 $m \leqslant n$，可能性是

$$P(\boldsymbol{X}=m \mid \boldsymbol{Y}=n)=\frac{1}{n}$$

因此，后验概率是①

$$P(\boldsymbol{Y}=n \mid \boldsymbol{X}=m) \propto \frac{1}{n^2}$$

当然，假定 $m \leqslant n$。这样，我们现在计算

$$\begin{aligned} P(\boldsymbol{Y}>2m \mid \boldsymbol{Y} \geqslant m) &= \sum_{n=2m+1}^{\infty} \frac{1}{n^2} \Big/ \sum_{n=m}^{\infty} \frac{1}{n^2} \\ &= \left(\frac{\pi^2}{6}-\sum_{n=1}^{2m} \frac{1}{n^2}\right) \Big/ \left(\frac{\pi^2}{6}-\sum_{n=1}^{m-1} \frac{1}{n^2}\right) \end{aligned}$$

---

① 值得注意的是，这里是许多贝叶斯主义者接受的某些东西。先验概率是不恰当的，但后验概率分布被认为是先验概率与可能性的乘积，是恰当的，这是杰弗里斯关注的问题所在。

因为 $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$。这个概率很接近于 1/2,因此,概率

$$P(\mathbf{Y} \leqslant 2m \mid \mathbf{Y} \geqslant m) \tag{1}$$

接近于 1/2。特别是,如果 $m = 50$, (1) = 0.507 5;如果 $m = 100$, (1) = 0.503 7;如果 $m = 1\,000$, (1) = 0.500 4;如果 $m = 10.000$, (1) = 0.500 0。 190

正是杰弗里斯相信这样的客观先验概率,才把他与第 7 节讨论的类似于德·菲内蒂(de Finetti)的最重要的主观主义者区分开来,而且,正如我已经评论的那样,这也是为什么我会在本节包括了对他的工作的关注。大多数主观主义者坚持认为,杰弗里斯的先验概率是不切实际的,因为没有关于问题特征的任何信息,每一个人都知道,在世界上的任何一个城镇里的电车数,在任何一种条件下,都不可能超过十万辆。但当然,把 n 的最大值固定在十万不会在很大程度上改变杰弗里斯对后验概率的分析。关于这种分析带来的其他问题不可能在这里讨论。

本节的其余部分专门讨论更明显的两个版本的概率的逻辑理论。第一个是卡尔纳普的确证理论,包括欣蒂卡(Hintikka)和后来的凯伯格(Kyburg)的推广在内,第二个是盖夫曼(Gaifman)、克劳斯(Krauss)和斯科特,还有丘瓦基(Chuaqui)的模型理论进路。

卡尔纳普的确证理论。关于这个主题的三个基本参考文献是由卡尔纳普(1945、1950、1952)提供的。伯克斯(Burks)、凯梅尼(Kemeny) 内格尔和普特南对卡尔纳普思想的出色的批评和卡尔纳普的答复,会在施利普(Schlipp)的著作中(1963)找到。这个理论的基本思想是,相对于一组特定的证据,提出一个假设的确证度概念。例如,通过观察六头黑牛和一只黑羊组成

的证据,给出所有的羊都是黑的这个假设的确证度有多大呢?从一种概率的观点来看,我们预期,确证度概念具有条件概率的形式特性。

遵循卡尔纳普等人的传统,我们将首先运用一个明显的语言框架。(然而,这是不必要的,正如我们看到的那样,我们完全能通过考虑集合论的实体来进行。)我们考虑由下列要素组成的语言:

(i) 圆括号,

(ii) 语句连接词,

(iii) 有穷个谓词数,

(iv) 个体的名称,

(v) 取值最多是可数个体数的个体变量,

(vi) 运用个体变量的量词,

(vii) 恒等号。

暂时,我们限于一阶谓词。因此,在这个框架内:

定义 1. *L_N^m 是满足(i)到(vii)的一种语言,具有 m 个一阶谓词和 N 个个体的名称。*

既然最初我们关注任意一种语言 L_N^m 的基本概念,我们将在定义和定理中省略对 L_N^m 的形式参考。我们也要求,一种语言 L_N^m 满足两个必要条件或约定:

191 **约定 1.** 不同个体的名称指不同的个体。

约定 2. 所有的谓词(在一个给定的语言中)在逻辑上都是独立的。例如,在相同的语言中,我们不允许谓词“是有色的”和谓词“是绿色的”,因为任何一种绿色的东西都是有色的。

我们接受在 2.2 节中已经定义的形式合理的分子式的一般概念,但某些进一步的特殊概念是有用的。

定义 2. *一个句子 s 是原子的,当且仅当,s 是由一个谓词*

和一个个体的名称组成的一个形式合理的分子式。

这里和后面为了达到举例说明的目的，让我们考虑一种特殊语言 L_3^2，它的谓词是"*R*"("是红的")和"*G*"("是光滑的")。三个个体名称是"*a*"、"*b*"和"*c*"。原子语句的例子包括"*Ra*"和"*Gc*"。

定义 3. **一个语句 *s* 是一种状态描述，当且仅当，*s* 是由一个词典顺序的原子语句的联合和原子语句的否定组成的一个形式合理的分子式，使得在 *s* 中每个个体的名称后面的每个谓词恰好出现一次。**

在 L_3^2 中状态描述的两个例子如下：

$$s_1 = \text{'}Ga \& Gb \& -Gc \& Ra \& -Rb \& -Rc\text{'}$$

$$s_2 = \text{'}-Ga \& -Gb \& -Gc \& -Ra \& -Rb \& Rc\text{'}$$

我们知道，使定义 3 的联合成为惟一的，我们必须固定某些谓词和个体名称的词典顺序的索引。在 L_3^2 中状态描述的总数是 $((2)^2)^3$ 或 64。因此，一般情况下，在任何一种语言 L_N^m 中，状态描述数是 $2^{m \cdot N}$。从直观上看，每一种状态描述都对应一个可能世界；不同的状态描述对应不同的可能世界，或者，用标准的概率理论的语言说(参见第 1 节)，对应不同的可能结果。我们现在转向某些标准的逻辑概念的明显的定义。

定义 4. **一个语句 *s* 的范围是这个语句成立的所有状态描述的集合。也就是说，如果一种状态描述在 *s* 的范围内是真的，*s* 就必须是真的。**

定义 5. **一个语句是 *L* -真的，如果它的范围是所有状态描述的集合。**

定义 6. **一个语句是 *L* -假的，如果它的范围是空集。**

例如，"$Ra \& -Ra$"是 *L* -假的。

定义 7. 语句 s_1，L-蕴含语句 s_2，当且仅当，s_2 的范围包括了 s_1 的范围，也就是说，$R(s_1) \subseteq R(s_2)$。

定义 8. X 是所有语句 s 的集合。

我们接下来在一种语言 L_N^m 的语句集合上定义一个正则测度的重要概念(卡尔纳普的术语)。这样一种测度只是在这种语言的状态描述集上的一个离散的概率密度，通过明显的相加条件扩展到这种语言的其他语句。

定义 9. 一个实值函数 m 是一个正则测度函数，当且仅当，

192 (i) m 的定义域是 X；

(ii) 如果 s 是一种状态描述，那么 $m(s) \geqslant 0$；

(iii) $\sum_{s \in S} m(s) = 1$，这里，S 是状态描述集；

(iv) 如果 s 是 L-假的，那么 $m(s) = 0$；

(v) 如果 s_1 不是 L-假的，而且 $R(s_1)$ 是 s_1 的值域，那么 $m(s_1) = \sum_{s \in R(s_1)} m(s)$。

对概率的逻辑理论来说，作为对逻辑演绎概念的扩展，我们需要提出一个与证据和假设相关的稍微弱一些的概念。证据和假设很少通过一个演绎推理系统相联系。因此，在实践中，我们将需要一种方法来表明证据与假设之间固定关系。我们能研究一个比较的确证概念：

$$C(h, e; h', e')$$

当 e 确证 h 大于 e' 确证 h' 时，这种关系成立。然而，取而代之，我们从一个定量的测度函数开始。中心问题是，选择什么样的正则测度函数。在无数的正则测度函数当中，哪一个最好地近似于我们先于任何信息的合理分布的直观思想呢？我们能够定性地接近于这类可行的测度。例如，我们试图提出会产生一个正则测度函数的定性公设，但稍后为了概率的其他概念，将会详

细地阐述这条进路。按照卡尔纳普的思路,我们将反而把先天条件直接强加于这些可能的测度。

首先,我们注意到,根据现有的一个测度函数,我们可以直接定义一个确证函数

$$c(h,\ e) = \frac{m(h \& e)}{m(e)}$$

假定 $m(e) > 0$。如果假设 e,L-蕴含 h。那么从我们的测度概念来看,

$$m(h \ \& \ e) = m(e)$$

因此,在这种情况下确证函数等于1,即 $c(h,\ e) = 1$。另一方面,如果 e,L-蕴含 $-h$,那么

$$m(h \ \& \ e) = 0$$

而且,确证函数等于0,即 $c(h,\ e) = 0$。因为根据这个测度函数,这个确证函数是立即可定义的,所以,我们的整个分析将关心什么样的特性对于一个测度函数的要求是合理的。从卡尔纳普的观点来看,这不是一个经验问题,而是一个标准问题或逻辑问题。

我们首先引入**对称态**的测度函数概念。(用卡尔纳普的术语说,测度恰好是**对称的**,但我希望命名两种不同的对称性。)在直观意义上,这种思想是先天的,我们不应该以任何一种方式在个体之间作出区分。例如,在 L_3^2 中,一个对称的测度函数 m 将有下列特性:

$$m(\text{`}Ra\text{'}) = m(\text{`}Rb\text{'}) \tag{2}$$

和

193

$$m(\text{`}-Ra \ \& \ Gb\text{'}) = m(\text{`}-Rc \ \& \ Ga\text{'}) \tag{3}$$

如果我们接受一个语句并考虑个体名称的任何一种排列，这个语句的一种对称测度仍然是相同的。在上面的(3)中，我们用到的排列是，$a \to c$，$b \to a$ 和 $c \to b$。这直观上对应于改变了演绎推理中的变量。这个概念本身没有无差异原理(贝叶斯的公设)有说服力，因为它没有确定关于状态描述的惟一的先验概率分布。

定义 10. 两个状态描述 s_1 和 s_2 是同构的，当且仅当，通过个体名称的一种排列，可以从 s_2 中获得 s_1。

在这种语境中，一种排列是把一组名称映射到自身的一个一一对应的函数。

定义 11. 在 X 上的一个正则测度函数 m 是对称态，当且仅当，m 把相同的测度赋予同构的状态描述。

有时有人建议说，我们超越了状态对称并选择一个先验测度作为惟一的正则对称态的测度函数，这个测度函数把相同的测度赋予所有的状态描述。我把这种测度函数称为 m_w，因为它是维特根斯坦在他的论著《逻辑哲学论》(1922)一书中有点晦涩地提出的一种建议。(参见 * 4.4、* 4.26、* 5.101 和 * 5.15；也参见卡尔纳普 1945，第 80—81 页。)

我能举一个例子从直观上确定，基于 m_w 的确证函数的不充分性。考虑语言 L^1_{101}。状态描述的个数显然是 2^{101}。对于任何一种状态描述 s 而言，通过下列等式给出测度：

$$m_w(s) = \left(\frac{1}{2}\right)^{101}$$

现在，设证据 e 是"$Pa_1 \ \& \ Pa_2 \ \& \ \cdots \ \& \ Pa_{100}$"，假设 h 是"Pa_{101}"。于是，$h \ \& \ e$ 是一个状态描述，因此，

$$m_w(h \ \& \ e) = \left(\frac{1}{2}\right)^{101}$$

当然,e只在两种状态描述中成立,即 $h \& e$ 和 $-h \& e$。因此,

$$m_w(e) = \left(\frac{1}{2}\right)^{101}$$

这样,

$$c_w(h, e) = \frac{m_w(h \& e)}{m_w(e)} = \frac{1}{2}$$

由于否认"P"在 e 中发生 m 次(其中,$1 \leqslant m \leqslant 100$),由 e 产生 e',如果用 e'取代 e,我们仍得到 $c_w(h, e') = \frac{1}{2}$,这是很不能令人满意的。

我们现在转向卡尔纳普的结构描述的重要概念。运用这个概念是他的确证理论的最显著的特征。

定义 12. 设 s 是一种状态描述。于是,通过 s 产生的结构描述是这个语句:它同构于 s 的所有状态描述的词典顺序的析取。

在非正式的意义上,我们通常会发现,把结构描述看成是状态描述的集合,是方便的。 194

例如,考虑语言 L_2^1。状态描述 s_i 和结构描述 S_j 是:

$$s_1 = \text{`}Ra \& Rb\text{'}$$

$$s_2 = \text{`}Ra \& -Rb\text{'}$$

$$s_3 = \text{`}-Ra \& Rb\text{'}$$

$$s_4 = \text{`}-Ra \& -Rb\text{'}$$

$$S_1 = \{s_1\}$$

$$S_2 = \{s_2, s_3\}$$

$$S_3 = \{s_4\}。$$

我们现在定义对称结构的测度。

定义 13. 在 X 上的正则测度函数 m 是对称结构,当且仅当,m 把相同的测度赋予所有的结构描述。

卡尔纳普没有这样阐述问题,下面这点是我们在这种发展中给出的那些定义的一个直接结果,即我们有下列表征定理,这个表征定理的证明是明显的。

定理 1. (卡尔纳普的表征定理)。对于任何一种语言 L_N^m 来说,存在着一个惟一的正则测度,它既是对称态,也是对称结构。

这种惟一的测度被卡尔纳普用 m^* 来表示。卡尔纳普没有使用**对称结构**的语言,但似乎很自然的是,用这个术语引入了两种明显的对称性。

在语言 L_2^1 的一个例子中,测度 m^* 是:

$$m^*(S_1) = m^*(S_2) = m^*(S_3) = \frac{1}{3}$$

由此得出

$$m^*(s_1) = m^*(s_4) = \frac{1}{3}$$

$$m^*(s_2) = m^*(s_3) = \frac{1}{6}$$

因为 m^* 也是对称态。

下列定理适用于具有多元谓词的语言,注意,这个定理独立于精确的谓词数,结果只依赖于个体名称数。后面,我们将用这个定理来确定 m^* 的一个不合需要的特性。

定理 2. 在由有限个谓词(不需要是一元的)和 N 个个体名称组成的语言中,一种结构描述中的状态描述的个数是一个 $N!$ 的除数。

例如，在 $N=5$ 和一个一元谓词的情况下，我们在任何一种给定的结构描述中有作为可能的状态描述数是：1、2、3、4、5、6、8、10、12、15、20、24、30、40、60 和 120。

定理 2 的证明包括了这里将不说明的关于有穷群的某些简单事实。建议没有群论知识的读者略去这个证明。 195

证明：设 S 是（满足我们假设的语言 L_N 的）一种结构描述，设 s 是在 S 中的任意一种状态描述。设 S_N 是 N 个字母的对称群，设 π 是在 S_N 中的一个排列，再设 $\pi^* s$ 是通过把排列 π 应用于（L_N 的）个体名称从 s 得出的状态描述。很容易证明，对于 S_N 中的每一个 π，$\pi^* s$ 在 S 中，同样，对于 S_N 中的 π、γ，

$$\pi^*(\gamma^* s) = \pi(\gamma)^* s$$

我们把排列集 S 定义为

$$S = \{\pi: \pi \in S_N \ \&\ \pi^* S = S\}$$

如果 π，$\gamma \in S$，那么

$$s = \pi^* s = \gamma^* s = \pi^{-1}\gamma^* s$$

因此，显然，S 是 S_N 的一个子群。现在，如果 $\sigma \in S_N$ ①，那么，$\sigma G(s)$ 是 G 在 S_N 中的一个左陪集，使得 $\rho \in \sigma G(s)$，当且仅当，在 G 中存在着一个 π，使得 $\rho = \sigma\pi$。

我们容易获得，如果 ρ 和 σ 在 S_N 中，那么

$$\rho^* s = \sigma^* s \Big|_{\text{当且仅当}\rho \in \sigma G(s)}$$

因为如果 $\rho^* s = \sigma^* s$，那么 $(\sigma^{-1}\rho)^* s = s$，$\sigma^{-1}\rho \in G(s)$ 以及 $\rho \in \sigma G(s)$；而且，如果 $\rho \in \sigma G(s)$，那么在 G 中有一个 π，使得 $\rho =$

① 原文 $\sigma e S_N$ 是印刷错误，与作者沟通后改正。——译者

$\sigma\pi$，由此，

$$\rho^* s = \sigma\pi^* s = \sigma^* \pi^* s = \sigma^* s$$

所以，

$$\rho^* s \neq \sigma^* s, \text{当且仅当}, \rho \notin \sigma G(s)$$

而且，在 S 中不同的状态描述数只是在 S_N 中 S 的左陪集数，即是 G 在 S_N 中的指数。但是，$G(s)$的指数是一个 $N!$ 的除数，即 S_N 的次序的一个除数。

证毕。

我们将用对确证理论进行这种带有一些批评和评论的分析作为结论，是为了从特殊发展到一般。首先，我们将考虑测度函数 m^*；接着考虑状态描述概念，然后考虑更一般的主题。

Ⅰ. 这种语言谓词在某种强的意义上必须是完备的，因为可以通过把既不在 h 中出现也不在 e 中出现的一阶谓词增加到这种语言中，来改变 $c^*(h, e)$。这种结果完全是反直觉的。

Ⅱ. 在 L_∞^m 中，在作为一个有穷样本的一种结果的任何一个普遍规律的 c^* 条件下，这种确证是零。(这里，L_∞ 有一个不同个体名称的无穷序列。)关于不同的相关结果，参见费金(Fagin, 1976)。

Ⅲ. 设 h 是 L_N^m 的一个语句。于是，如果对于每个 N 来说，h 在 L_N^m 中测度等于 h 在 L_{N+1}^m 中的测度，那么 m^* 就是适当的。卡尔纳普证明，当我们限于一元谓词时，m^* 是适当的。这在直觉上是相当合乎需要的特性，因为证据 e 对 h 的确证无疑不应该随着既不在 h 中出现也不在 e 中出现的新个体名称的引入而发生变化。下列反例表明，对于含有二元谓词的一种语言来说，m^* 是不适当的(Rubin and Suppes, 1995)。

196

设 W 是一个二元谓词。设“a”和“b”是含有一元谓词“W”

的 L_2^1 的个体名称。设 i 是语句

$$\text{“}Waa \ \& \ Wab \ \& \ Wba \ \& \ Wbb\text{”}$$

我们很容易证明,对于 L_2^1 来说,$m^*(i)=0.1$。现在,设 $s(N)$是在 L_N^1 中的结构描述数。从定理 2 立即可以得出,任何一个语句 i 的测度必须等于

$$m_N^*(i)=\frac{x}{s(N)\cdot N!}$$

其中,x 是正整数。在 m_3^* 的情况下,我们有

$$m_3^*(i)=\frac{x}{104\cdot 3!}$$

并且,显然不存在正整数 x 使得

$$x=(.1)\cdot(104)\cdot(3!)$$

事实上,

$$m_3^*(i)=\frac{23}{312}$$

运用戴维斯(R. L. Davis)的结果(1953),我们有

$$s(4)=3\,044$$

显然,没有一个整数 x 使得

$$\frac{x}{(3\,044)4!}=\frac{23}{312}$$

结果

$$m_3^*(i)\neq m_4^*(i)$$

Ⅳ. 要求一元谓词的逻辑独立性是不现实的和不能实现的。甚至更糟糕的是,允许 n 元谓词没有形式特性,是不合需要

的。因此，通常假定"比……暖和"具有作为逻辑特性而不是事实特性的非自反性的特性。这个问题在文献中一直有广泛的讨论。

Ⅴ. 所有的表达根据基本谓词必须是可定义的，但不清楚如何给出一个充分而完备的基本谓词集。一个相关的和同样基础的问题是，像平常在经验科学中运用的那样，引入数值函数。

Ⅵ. 我前面提到过，在一种形式语言的语境中，真的根本没有必要提出确证理论。在卡尔纳普的进展中，没有以任何一种基本的方式使用这种语言的系统的元数学特性，而且，我们通过谈论第 1 节和第 2 节定义的那类有穷概率空间，恰好可以很好地继续下去。此外，在那种语境中，我们能够以一种稍微不同的视角表达对测度函数的选择，以及表达这种选择与我们前面在应用概率的经典定义时关于对称性原理的讨论的关系。

在处理含有两个一元谓词和两位个体的一种语言中，让我197 们首先接受对确证理论的集合论表述。一种状态描述的概念只能被一种可能结果的标准概念所替代，而且，其他定义根据逻辑概念和集合概念之间的标准关系来进行。例如，谈论一个语句 L蕴含着另一个语句，被谈论一个事件蕴含着另一个事件所替代，这被解释为只意味着第一个事件是第二个事件的一个子集。

在这种集合论框架内，我们如何把卡尔纳普对 m^* 的特别选择表示为一个首选的测度函数是显然的。正是这种测度函数源于应用两个对称性原理。在根据基本特性的有穷集定义可能结果的情况下，一种结构描述现在只被定义为，从通过一种个体排列的已知结果获得的所有可能结果的集合——恰好是前面给出的状态对称性的定义。另一个原理是所定义的结构对称性原理。

可能结果的集合的这种先天结构似乎不是自然的，这方面

的一个简单例子是，把一枚硬币掷 N 次的试验。这些可能结果恰好是反映正反面的每一个可能序列的长度 N 的有穷序列。一种结构描述是有相同的正面数的可能结果的集合。如果我们把卡尔纳普的测度 m^* 应用于这种情形，我们赋予掷出全是正面的可能结果的权重，比掷出正面的个数例如是近似于 $N/2$ 的权重要大得多。对这个例子而言，事实上，我们更喜欢维特根斯坦的测度 m_w，而不喜欢卡尔纳普的测度 m^*。换言之，在这个例子中，卡尔纳普的对称性原理似乎几乎没有被看成是自然的和适当的。

此外，在许多实验的设计中，不管可能结果的集合是有穷的还是无穷的，应用对称性原理，迫使以一种相当有效的方式选择一种先验分布。例如，在处理反应的一个闭联集的实验中，在受试者中间，任意地随机打乱根据左-右和上-下给这个闭联集的部分所贴的标签。这是以这样一种方式进行的：不管受试者最初的偏见是什么，在最初的反应分布中必定满足相当有说服力的简单的对称性原理。当然，这样的一个例子不是卡尔纳普自己在讨论选择一个测度函数时所正视的那类问题，这是真的；但在我自己的判断中，这两者之间的关系要比他可能愿意承认的关系更紧密。

Ⅶ. 卡尔纳普(1950，第 19—51 页)特别强调指出，在他的观点中，有两个不同的概率概念。一个概念是经验的相对频率概念，另一个是在确证理论中表达的概率概念。他最强有力地强调说，在特征上，统计概念是经验的，而确证理论概念是分析的或逻辑的。尽管这种区分似乎已经被极其大量的哲学家所接受，但是，卡尔纳普根本没有为这种二元论的进路明确地举出一个有说服力的事例。

我肯定承认，在所有的理论科学中，理论几乎总是有两个密

切相关的模型。一个模型是作为纯数学的模型提供的，另一个被看成是以某种直接的方式从经验数据中推论出来的模型。然而，这两种模型的存在并没有证明卡尔纳普极力主张的两个概率概念之间的区分是合理的。正如卡尔纳普所建议的那样，在理论决定的概率和经验决定的概率之间有一个自然的二分法，但这种二分法的要点根本不是卡尔纳普建议的要点。在确证理论与统计理论之间没有可比性，但理论模型和实验数据之间只是普通的比较。

198

看待 m^* 的另一种方式，也反对卡尔纳普的观点。这就是在贝叶斯意义上把 m^* 看成是一种客观的先验分布。于是，假设 h 在证据 e 上的确证度 $c^*(h, e)$ 恰好是贝叶斯意义上的后验概率。基于这种观点，肯定不存在两种概率：逻辑概率和统计概率，统计概率被用作补充 m^* 和 c^*，或者给定 h 后 e 的相关可能性。正如某些贝叶斯主义者喜欢说的那样，如果使用从对称性原理推出的一个客观的先验概率，那么统计学在特征上就是纯数学的或演绎的。证据 e 当然是经验的，但是，有人坚持认为，已知 h, e 的可能性不是经验的。换一种角度来看，我们能够为任何一个可能的证据 e 计算 $c^*(h, e)$。这种计算不是经验的，也根本没有加入概率的统计意义，只有什么是实际证据的非统计的经验问题。我强调，这里我只触及重要而敏感的论题，而且，不可能在本章的框架内充分详细地考察这些论题。然而，我希望，我足以说出了对卡尔纳普的二元的概率概念的有效性产生的怀疑。①

---

① 在卡尔纳普关于确证理论的长期工作之后的这些年里，关于最近对确证理论的一种极好的一般评论，参见菲特尔森(Fitelson, 2001)，在这个文献中，充满了反对各种不同提议的精彩反例和论证。

欣蒂卡的两个参量的理论。已知在为一种惟一的选择(比如 m^*)进行辩护时有许多困难,卡尔纳普退回到用含有一个自由参量 λ 的一种测度 m_λ(Carnap, 1952)。这种新的测度,像 m^* 一样,具有对称态的特性。似乎显然的是,除非能够给出特殊的自变量,来选择 λ 的一些特殊值,因而选择不同的归纳逻辑,否则,就几乎没有理由把这种新理论看成是一种纯逻辑的概率理论。不可避免的是,选择 λ 的一个特殊值来惟一地确定一种测度,似乎就在概率的主观理论的重要传统之中。

但是,我认为,更有用的是考察后来由欣蒂卡概括出来的两个参量的理论(1966),而不是探索卡尔纳普的归纳逻辑的 λ 理论,欣蒂卡也提供了他的体系与卡尔纳普理论之间的一种极好的详细比较。

欣蒂卡引入两个参量 α 和 λ,这个 λ 与卡尔纳普的 λ 基本相同。倘若有穷的证据完全是正的,欣蒂卡的第二个参量 α 的主要理论依据是提供一个框架,在这个框架内,逻辑上不真的语句仍然可以在个体的一个无穷全域中得到一个非零的确证度,即正的条件概率。大体上说,卡尔纳普的 λ 和欣蒂卡的 λ 是通过由各种对称性来确定先验概率的已知权重的一个指数。λ 越大,这些先验考虑的权重就越大。

尤其是为了处理一个无穷全域的渐近情况,从状态描述转到较少依赖于特殊个体的框架,是十分必要的。假设语言是 L_N^m,即存在着 m 个基本的一元谓词 $P_1 \cdots P_m$。卡尔纳普为这些基本谓词或它们的否定的合取引入 Q-谓词这个术语。因此,当存在着 m 个基本谓词时,存在着 2^m 个 Q-谓词。在定义 2 之后考虑的 L_3^2 的例子中,有四个 Q-谓词: 199

$$Gx \; \& \; Rx,$$

$$Gx \ \& \ -Rx,$$

$$-Gx \ \& \ Rx,$$

$$-Gx \ \& \ -Rx。$$

用欣蒂卡的术语说，一种构成是在 L_N^m 中的一个语句，这个语句准确地说出哪些 Q-谓词是实体化的。如果这样的个数是 w，欣蒂卡（含糊地但有用地）称之为构成 C_w。在上面的 L_3^2 的例子中，让我们根据已知顺序把四个 Q-谓词缩写为：$Q_1(x)$，…，$Q_4(x)$。于是，样本"$Ga \ \& \ -Ra$"、"$-Gb \ \& \ Rb$"和"$Gc \ \& \ -Rc$"将例示了下列构成：

$$(\exists x)Q_1x \ \& \ (\exists x)Q_2x \ \& \ (\forall x)(Q_1x \vee Q_2x)$$

欣蒂卡把他的理论集中于先验概率 $P(C_w)$ 和可能性 $P(e|C_w)$，然后通过贝叶斯定理（第 1 节的定理 8）计算后验概率 $P(C_w|e)$。我这里限于考虑先验概率 $P(C_w)$。

为了给出对 α 和 λ 的渐近值（其值域是从 0 到∞）的一种看法，下面我们有：

(i) 当 $\alpha = \infty$ 和卡尔纳普的 Q-谓词数 $\lambda = 2^m$ 时，我们得到卡尔纳普的 m^* 。

(ii) 当 $\alpha = \infty$ 和 $\lambda = \infty$ 时，我们得到维特根斯坦的 m_w。

(iii) 当 $\alpha = \infty$ 时，我们得到卡尔纳普的 λ-闭联集。

欣蒂卡以卡尔纳普的 λ-体系为出发点。已知一个固定的构成 C_w，α 个个体与它相容，即例示了 C_w 的 Q-谓词的先验概率是：

$$\frac{b}{2} \cdot \frac{1+b}{1+\lambda} \cdot \frac{2+b}{2+\lambda} \cdots \frac{\alpha-1+b}{\alpha-1+\lambda} \tag{4}$$

其中 $b = \dfrac{w\lambda}{2^m}$。在欣蒂卡的符号中，(3)的分子是 $\pi(\alpha, b)$，分母是

$\pi(\alpha, \lambda)$。他设置构成的先验概率正比于(3)。因此,

$$P(C_w) = \frac{\pi(\alpha, b)}{\sum_{i=0}^{K} \binom{K}{i} \pi\left(\alpha, \frac{i\lambda}{K}\right)} \quad (5)$$

其中 $K = 2^m$。显然,把相同的先验概率赋予例示了相同 Q-谓词数的两种构成的对称性是这种设置的组成部分。同样显而易见的是,我们不可能从一般的对称性原理推出像(4)一样的等式,事实上依赖于两个非负的实参量 α 和 λ。相应地,我们不能指望根据定理 1 的精神拥有一个表征定理。

欣蒂卡的系统有许多好的特征,但似乎对我来说,应该主要把它看成是对主观理论的贡献。对两个自由参量值的选择,被 200 预计为是因人而异的,也是随情境而变化的。这恰好是纯贝叶斯式的那类问题,对主观观点的最强有力的辩护,希望是这种情况。

凯伯格。凯伯格(Kyburg)已经提出了一个更丰满和更现实的逻辑理论(1974)。这仍然属于卡尔纳普在他的早期工作中如此强调的明显的逻辑语言传统。这里是为什么凯伯格的表述比卡尔纳普开始的表述更丰满和更现实的原因所在。首先,他引入了一种更丰满的形式语言,能够非常适当地以一种形式的方式表示标准的统计概念,比如,随机抽样和检验假设。第二,因为这种更丰满的语言,能够分析来自统计实践——至少在统计学教学中的课堂实践——的那类更现实的事例和问题。第三,表述如此发展的理论并非要求一个已知问题或实验的一种先验概率或分布,而是与一个分布族一起很好地起作用。

凯伯格对概率基础有许多令人感兴趣的评论。他的令人满意的论述应该比过去几十年产生的影响更大。我不能确定为什么并非如此,但我猜测,主要是因为它过分地使用了逻辑符号和

形式逻辑的概念,比如,把随机性和概率定义为是元语言的概念。绝大多数的统计学家和数学家很不喜欢这种过分形式化的进路。卡尔纳普也没有对这批人产生很大的影响,但凯伯格的进路当然是一个好消息,因为他为研究统计学和概率的许多标准困惑作出了细致的努力,正如独立于任何一种明显的逻辑形式语言所阐述的那样。

模型论进路。给弗盟(1964)、斯科特和克劳斯(1966)、克劳斯(1969),还有盖夫曼和斯尼尔(Snir)(1982)一直关注为一般的一阶语言提出一个概率测度的概念,还借助拥有一个 σ-加法理论,把这种进展推广到非限定性语言。这种工作的精神不是通过推测性的逻辑论证来证明选择惟一的先验测度是合理的。更确切地说,其意图是把普通逻辑推广到*概率逻辑*。斯科特和克劳斯的核心概念是在有概率 α 的一个概率系统 Ω 中成立的一个语句的概念。这个概念类似于塔斯基的满足普通逻辑的概念。从这个概念和相关概念出发,作者建构了一个漂亮的形式理论。

然而,实际上,人们没有把这种模型理论的进路当作是对在杰弗里斯或卡尔纳普的意义上的概率的逻辑基础的一种贡献。事实上,盖夫曼、斯科特以及克劳斯的大多数结果类似于从柯尔莫哥洛夫的公理中推出的标准概率理论中的结果,它们对于基础基本上是中性的。在这一节的意义上,他们的模型理论进路没有为概率基础提供一个新的表征理论。

丘瓦基。一个有意义的表征定理是由罗兰多·丘瓦基(Rolando Chuaqui)运用模型理论的进路和概率的语义定义提供的(1977,1991)。但是,他证明的这个表征定理实际上不依赖于他的模型理论方法,而是依赖于把这些方法推广到非标准的无穷小分析。这里,为无穷小分析引入适当的形式概念和符号,

201

在技术上太离题了,但我相信,为丘瓦基的表征结果提供一个清楚的直观描述是有可能的。一个**标准的**随机过程恰好是被一个离散的或连续的集合 T 指示的随机变量族 $\{\mathbf{X}_t: t \in T\}$。一个**非标准**的这种过程也是由无穷小或它们的倒数,即超实数指示的。在纳尔逊(Nelson)的著作中(1987)最明确的观点是,我们通过运用非标准的无穷小方法,能够极大地简化受到在科学上不重要的但复杂的数学技巧所困扰的概率的许多部分。例如,很容易定义一个时间连续的离散态的随机过程,这个过程能够在有穷时间内经历无穷个状态数。更加糟糕的是永久的测度理论问题,这些问题困扰着连续时间、连续状态的随机过程。纳尔逊针对这些问题的一个更简单的理论,提供了在所有的普通科学语境中都可运用的重要的指导原则。

在纳尔逊的框架内,通常的时间连续的过程被一个超有穷的过程所替代。例如,设索引集合 T 是自然数的集合。它能够被$\{0, 1, \cdots, \nu\}$所替代,其中,ν 是一个无穷大的自然数。(如果 $1/\nu$ 是无穷小,比如说 ϵ,而且,对于所有标准的有穷实数 x 而言,ϵ 有特性 $0 < \epsilon < x$,那么,自然数 ν 就是无穷大。)第二个例子是,设 T 是闭区间$[0, 1]$。我们定义离散时间集合 $T' = \{0, dt, 2dt, \cdots, 1\}$,其中,对于某个无穷大的自然数 ν 来说,$dt = \frac{1}{\nu}$。因此,T'是超有穷的。T'被说成是 $T = [0, 1]$ 的近区间。纳尔逊(1987)证明,在所有概率量的无穷小等价的范围内,把任何一个标准的随机过程近似地表示为被一个邻近索引集所索引的一个邻近过程。对于随机过程来说,这种非形式地陈述的定理本身是一个有意义的形式表征定理。

丘瓦基用它证明下列表征定理(我在这里陈述了他的基本定理的推论,Chuaqui 1991,p. 351)。

定理 3. (Chuaqui 1991)。对于被一个超有穷集合 T 所索引的任何一个随机过程来说,在它的所有元素都是同样可能的概率空间上,存在着一个接近无穷小等价的随机过程。

正如丘瓦基所注意到的那样,运用上面非形式地陈述的纳尔逊的定理,就很容易地把他的定理推广到我所说的任何一个标准随机过程的拉普拉斯表征——之所以说是拉普拉斯的表征,是因为这个表征概率空间满足推扩了的拉普拉斯的第一概率原理,即在概率集合上的一种均匀分布。

这个表征定理没有看起来那么令人惊奇。在第 1 节引入随机变量时已经强调过,在其上定义一个随机变量的概率空间,不能通过这个随机变量的分布(主要感兴趣的目标)惟一地加以确定。因此,对于有穷概率空间来说,为了在这个有穷空间上定义任何一个随机变量,不难找到一个等概率原子事件的等价空间。非标准的无穷小分析允许把这些思想直觉地推广到含有超有穷索引集合的任何一个随机过程。

5.6 概率的倾向性表征[①]

202 近年来,概率的倾向性解释(最初被认为是一种客观解释)深受许多哲学家的欢迎,而且,关于这个问题的现有哲学文献,范围极其广泛。倾向性概念本身不是一个新的概念。《牛津英语词典》引用了倾向性的明确而简单的用法,它的一般意义在 17 世纪已经出现,例如,源于 1660 年“这些植物为什么有……生根的倾向?”[②]这种概率解释的思想运用了具有不同物理倾向

---

① 本节大量利用了苏佩斯(1987)的研究。

② 此外,在休谟的《人性论》(1739)中有许多用法。

的对象的一般思想，例如，放在水中会溶解的倾向，然后，把这种思想推广到概率。从这些讨论也清楚地看出，倾向性可能被看成是一种习性。[这方面的一个相当详细的讨论能够在梅勒(Mellor)著作的第四章找到(1971)。]

对概率的倾向性解释的最卓越的提倡者是波普尔(Popper)，他在两篇有影响的文章中提出了主要思想(1957,1959)。波普尔作为他发展倾向性解释的主要动机之一是提出，需要在量子力学中给出单一情形概率的一种客观解释，即个体事件概率的一种客观解释。正如在文献中所称的那样，单一情形的概率对主观主义者来说当然是没有问题的，但对相对频率的理论家来说，却是一个曲折的问题。关于我们如何思考既与单一情形的概率相关也与相对频率相关的倾向的一种非常详细的讨论，会在吉雷(Giere)的著作(1973)中找到。我同意吉雷的观点：离开相对频率理论的真正理由之一是单一情形的问题，因此，我们应该把惟一事件的倾向性解释看成是基本的或主要的。吉雷引用了许多原文表明，波普尔在这个问题上是犹豫不定的。

正如我在论文(1974a)中指出的那样，在关于倾向性解释的哲学和科学基础的这些极好的直观讨论中所缺少的是，有必要对倾向性解释证明一些东西。在我建议的框架内，这相当于是证明一个表征定理。在波普尔对我的批评的回应中(1974)，他提到他自己关于条件概率的思想和如何处理有零概率的证据问题。还有一些其他细节，但他没有抓住我批评的要点，由于他没有为倾向性提供不同的概念分析，所以，关于这种解释，真正有趣的东西是能够证明的。这样一种证明的要求不是无价值的，或者，不只是形式的要求。为了使概率解释成为有趣的，需要增加一些明确的概念，来超越被柯尔莫哥洛夫公理化的形式理论中的那些概念。我们能够直接把倾向性解释为概率，这似乎是

某些倡导者的一种强烈倾向，这种纯属劝告的评论，没有抓住提供一种更彻底分析的要领。

因为我认为，倾向性解释有许多好的方面，正如我在 1974 年关于波普尔的一篇文章中表明的那样，所以，我希望证明三个不同的表征定理，每一个表征定理都打算超越概率的形式理论，为倾向性提供一种分析，而且，作为第四个例子，也陈述了一个从经典力学来看是令人惊奇的定理。

203

第一个表征定理最接近概率本身，它应用于放射现象。第二个表征定理代表了作为反应强度实现的倾向性的心理现象。这种概率就是明确地从反应强度推出的。历史上最重要的第三个例子是，完全基于属于经典力学的考虑，由掷硬币的行为、轮盘赌和其他类似的方法派生出来的。第四个例子表明，纯粹的决定论系统如何能够产生出随机现象。

倾向性表征定理的这四个不同例子，在任何一种直接的意义上，都没有强迫在倾向性的单一情形的观点和相对频率的观点之间立即作出决定。但我当然看到，没有普遍的理由不运用这些定理来计算单一情形的概率。只要对一种已知的倾向性结构提出一种相对详细的说明，似乎就自然会这么做。

衰变的倾向。在转向技术发展之前，关于这里遵循的进路已经有一些一般的评论，这些评论多数取自苏佩斯的论文(1973a)。

第一种评论是关注这种事实：在所遵循的这些公理中，倾向性，作为定性地表达概率的一种手段，是事件的特性，而不是对象的特性。例如，把原始符号解释为断言，已知事件 B 的发生，事件 A 发生的倾向，至少与已知事件 D 的发生，事件 C 发生的倾向一样大。此外，事件 B 和 D 实际上可能并不发生。我们所估计的是，已知某个别的事件发生，这些事件发生的趋势或发

生的倾向。如果在谈论事件的发生时，运用倾向性的语言似乎不熟练或不自然，那么，为了谈论对象的特性，很容易简单地运用定性的概率语言和保留倾向性的语言。尽管我反对这种措施本身。无论如何，对于所提出的主要问题来说，这个论题是次要的。

第二种评论关注这里得到的这种表征定理和主观理论通常证明的那类定理之间的明显区别。正如在第 7 节考虑的那样，主观理论的典型特征是，引入强的结构公理，足以把惟一的概率测度强加给事件。这里所阐述的客观理论，正是缺少这种惟一性，但在我自己的判断中，缺乏惟一性是优势，而不是劣势。以放射性衰变为例。从概率的公理来看，如果没有正在衰变的物质的特殊实验或鉴别，我们就肯定不会预计，能够先验地得到衰变的几何分布的独特参量。正是这样一个参量的结果，即惟一性，形成了一个参量集，是客观理论的典型特征，我也主张，在从物理学到心理学的范围广泛的科学领域中，是标准实验的典型特征。换言之，概率的结构公理与必要公理一起，确定了概率测度的参量形式，但不是惟一地确定的。为了确定参量的数值，需要特殊的实验和基于经验数据估计参量的特殊方法。主观主义者通常以类似于当前通过假设参量空间上的一种概率分布的一个工作过程而告终。这些过程接近于这里所讨论的过程，因此，它们理应如此，因为人们就在实践中如何评价有点像一种几何分布或指数分布的一个参量，已经达成了很大的共识。 204

我想强调的重要观点是，在这个理论的基本立足点方面，主观主义者通常坚持惟一的概率测度，然而，似乎对我而言，对潜在的惟一的概率测度的这种承诺，对于大多数科学应用来说，是一个不切实际的前提。这不是说在主观主义的进路中存在着任何矛盾；只是说，从平常的科学实践的观点来看，当前的客观的

观点更加自然。[正如在第 3 节中已经注意到的那样,对倾向性表征的这种观点的支持可以在范·兰巴根(van Lambalgen, 1987a)的不同评论中找到。]

我也强调,客观理论的较弱的结构公理的直观意义,不同于主观理论的较强公理的直观意义。客观的结构公理被用来表示关于经验现象的特殊的定性假设或规律。这些公理的形式随着应用的不同而不同。一个特殊的结构公理提供了明确关注下面这点的方法,即把什么样的经验思想应用于一个已知的实验类。后面针对放射性衰变的特殊案例,更多地讨论这种观点。

我现在转向形式的发展。首先,在陈述必要的公理时,我将运用在概率的主观理论中是标准的这类定性的概率语言。真正出现倾向性的地方位于对下列问题的讨论中:即对反映了关于放射性衰变现象的强物理假设的相当特殊的结构公理进行讨论。

我不讨论下列定义中给出的公理,因为我在这里给出的公理,仅相当于是对克兰茨等人(1971,第 222 页)的定义 8 的前六个公理的小小修改,再加一个可替代的公理,即下面的公理 7。[①] 注意,$\approx$是指等价,即,

$$A \mid B \approx C \mid D \mid \text{,当且仅当,} A \mid B \geq C \mid D \text{ 和 } C \mid D \geq A \mid B$$

定义 1. 一个结构 $\Omega = (\Omega, \mathfrak{J}, \geq)$ 是一个定性的条件概率结构,当且仅当,$\mathfrak{J}$ 是在 Ω 上的集合的一个 σ-代数,而且,在有 $B, D, F, G \succ \varnothing$ 的 $\mathfrak{J}$ 中,对于每个 $A, B, C, D, E, F, G, A_i, B_i, i = 1, 2, \cdots$,下列公理都成立:

公理 1. 如果 $A \mid B \geq C \mid D$ 和 $C \mid D \geq E \mid F$,那么 $A \mid B \geq E \mid F$;

---

① 在第 7 节关于主观概率的定义 3 中,提供了相关的但更简单的充分必要的定性条件期望的公理。

公理 2. $A \mid B \succeq C \mid D$,或者 $C \mid D \succeq A \mid B$;

公理 3. $\Omega \succ \varnothing$;

公理 4. $\Omega \mid B \succeq A \mid D$;

公理 5. $A \cap B \mid B \approx A \mid B$;

公理 6. 如果对于 $i \neq j$, $A_i \cap A_j = B_i \cap B_j = \varnothing$,而且, 205
对于 $i = 1, 2, \cdots$, $A_i \mid D \succeq B_i \mid F$,那么 $\bigcup A_i \mid D \succeq \bigcup B_i \mid F$;此外,如果对于某个 i, $A_i \mid D \succ B_i \mid F$,那么 $\bigcup A_i \mid D \succ \bigcup B_i \mid F$;

公理 7. 如果 $A \subseteq B \subseteq D$, $E \subseteq F \subseteq G$, $A \mid B \succeq E \mid F$ 和 $B \mid D \succeq F \mid G$,那么 $A \mid D \succeq E \mid G$;此外,$A \mid D \succ E \mid G$,除非 $A \approx \varnothing$,或者 $A \mid B \approx E \mid F$ 和 $B \mid D \approx F \mid G$;

公理 8. 如果 $A \subseteq B \subseteq D$, $E \subseteq F \subseteq G$, $B \mid D \succeq E \mid F$ 和 $A \mid B \succeq F \mid G$,那么 $A \mid D \succeq E \mid G$;此外,如果任何一个假设都是 $\succ$,那么结论也是 $\succ$。

说定义 1 的公理是必要手段,当然,它们是下列假设的一个数学结果:在 $\mathfrak{I}$ 上定义一个标准的概率测度 P,使得

$$A \mid B \succeq C \mid D \mid \text{,当且仅当,} P(A \mid B) \geqslant P(C \mid D) \quad (1)$$

需要必要的公理和充分的结构公理共同确保存在满足(1)的一个概率测度,恰好这一点是依情况而定的。定义 1 的八个公理中的大多数公理很可能通常将是需要的。在许多情况下,也将要求一个阿基米德的公理;以几种形式之一对这个公理的阐述,在文献中是常见的。下面的版本来自克兰茨等人(1971,第 223 页)。

定义 2. 一个定性的条件概率结构 $\Omega = (\Omega, \mathfrak{I}, \succeq)$ 是阿基米德的,当且仅当,每个标准序列都是有穷的,其中,$(A_1, A_2, \cdots)$ 是一个标准序列,当且仅当,对于所有的 i, $A_i \succ \varnothing$,

$A_i \subseteq A_{i+1}$ 和 $\Omega \mid \Omega \succ A_i \mid A_{i+1} \approx A_1 \mid A_2$。

我现在转向一个定理,这个定理最早在苏佩斯(1973a)的定性概率的文献中似乎已经陈述过。前面没有出现的理由似乎是,当概率的考虑只限于特殊的物理情形时,在这种情形中,概率的主观理论经常使用的可解性公理通常不成立。结果,必须改变代表存在概率测度的熟悉的证明方法。没有这种改变,就不需要这个定理。

这个定理是关于标准序列的。只有这个定理不会产生概率测度,但它保证,存在着一个能够推广到所有事件的数值函数,因此,当表示结构的不必要约束的物理假设足够强时,它就成为一种测度。

定理 1. (标准序列的表征定理)。设 $(A_1, \cdots, A_n)$ 是一个有穷标准序列,即 $A_i \succ \varnothing$, $A_i \subseteq A_{i+1}$ 和 $\Omega \mid \Omega \succ A_i \mid A_{i+1} \approx A_1 \mid A_2$。于是,存在着一个数值函数 Q,使得

(i) $A_i \subseteq A_j$,当且仅当,$Q(A_i) \leqslant Q(A_j)$,

(ii) 如果 $A_i \subseteq A_j$ 和 $A_k \subseteq A_l$,那么,$A_i \mid A_j \approx A_k \mid A_l$,当且仅当,$\dfrac{Q(A_i)}{Q(A_j)} = \dfrac{Q(A_k)}{Q(A_l)}$。

此外,对于满足(i)和(ii)的任何一个 Q 来说,存在着一个 $q(0 < q < 1)$ 和一个 $c > 0$,使得

$$Q(A_i) = cq^{n+1-i}$$

206 **证明:** 设 $0 < q < 1$。把 $Q(A_i)$定义为

$$Q(A_i) = q^{n+1-i} \tag{2}$$

那么显然满足(i),因为一个标准序列的项是明确的;否则,将存在着一个 i,使得 $A_i = A_{i+1}$,因此,$A_i \mid A_{i+1} \approx \Omega \mid \Omega$, 与假设相反。这样,我们可以立即转向(ii)。首先,注意到下列等式:

$$A_i \mid A_{i+m} \approx A_j \mid A_{j+m} \tag{3}$$

通过归纳来证明(3)。对于 $m=1$,恰好是这种假设:对于每一个 i,$A_i \mid A_{i+1} \approx A_1 \mid A_2$。现在,假设它对于 $m-1$ 成立;我们就有

$$A_i \mid A_{i+(m-1)} \approx A_j \mid A_{j+(m-1)}$$

而且对于任何一个标准序列来说,

$$A_{i+(m-1)} \mid A_{i+m} \approx A_{j+(m-1)} \mid A_{j+m}$$

由此根据公理 7,像渴望的那样,$A_i \mid A_{i+m} \approx A_j \mid A_{j+m}$。接下来,我们表明,如果 $A_i \subseteq A_j$,$A_k \subseteq A_l$ 和 $A_i \mid A_j \approx A_k \mid A_l$,那么存在着一个 $m \geqslant 0$,使得,$j=i+m$ 和 $l=k+m$。既然 $A_i \subseteq A_j$ 和 $A_k \subseteq A_l$,所以,一定存在着非负的整数 m 和 m',使得 $j=i+m$ 和 $l=k+m'$。假设,$m \neq m'$,并且在不失一般性的情况下,假设 $m+h=m'(h>0)$。那么,显然,

$$A_i \mid A_{i+m} \approx A_k \mid A_{k+m}$$

此外,

$$A_{i+(m-1)} \mid A_{i+m} \approx A_{j+(m-1)} \mid A_{j+m}$$

同样还是根据公理 7,

$$A_i \mid A_{i+m} \succ A_k \mid A_{k+m'}$$

与我们的假设相反,因此我们必须有 $m=m'$。

根据这些结果,我们能够确定(ii)。作为一种条件,我们有 $A_i \subseteq A_j$ 和 $A_k \subseteq A_l$。首先假设 $A_i \mid A_j \approx A_k \mid A_l$。于是,我们知道,存在着一个 m,使得 $j=i+m$ 和 $l=k+m$,由此,

$$\frac{Q(A_i)}{Q(A_j)}=\frac{q^{n+1-i}}{q^{n+1-i-m}}=\frac{q^{n+1-k}}{q^{n+1-k-m}}=\frac{Q(A_k)}{Q(A_l)}。$$

第二,我们假设

$$\frac{Q(A_i)}{Q(A_j)} = \frac{Q(A_k)}{Q(A_l)}$$

根据 Q 的定义,一个简单的代数问题是看到,一定存在着一个 m',使得 $j = i + m'$ 和 $l = k + m'$,由此,根据我们前面的结果,$A_i \mid A_j \approx A_k \mid A_l$。

最后,像在这个定理中所表达那样,我们必须证明 q 的惟一性。假设存在着满足(i)和(ii)的一个 Q',使得根本不存在 $c > 0$ 和 $q(0 < q < 1)$,而且,对于所有的 i,

$$Q'(A_i) = cq^{n+1-i}$$

207 设

$$Q'(A_n) = q_1$$

$$Q'(A_{n-1}) = q_2$$

设

$$q = \frac{q_2}{q_1}$$

$$c = \frac{q_1^2}{q_2}$$

显然,

$$Q'(A_n) = cq$$

$$Q'(A_{n-1}) = cq^2$$

根据我们关于 Q'的假设,设 i 是最大的整数(当然 $i \leqslant n$),使得

$$Q'(A_i) \neq cq^{n+1-i}$$

我们有

$$A_i \mid A_{i+1} \approx A_{n-1} \mid A_n$$

由此通过(ii)，

$$\frac{Q'(A_i)}{cq^{n-1}} = \frac{cq^2}{cq}$$

因此，

$$Q'(A_i) = cq^{n+1-i}$$

与假设相反，并证明了这个定理。

我现在转向放射性衰变现象。放射性衰变现象是概率现象的一个最著名的物理学事例，对于这种概率现象来说，几乎没有人自称拥有更深层次的决定性理论。这里，为了简单起见，我将考虑导致衰变的概率的几何分布的一个离散时间版本。可直接推广到连续时间，但这里不考虑这个问题。在当前的语境中，重要的观点是概念的，而且，我希望使技术细节最小化。当然，在科学的其他领域，对衰变公理有完全不同的解释，其中的一些解释将在后面提到。

在概率理论的一种特殊应用中，第一步是描述样本空间，即可能实验结果的集合，或者作为一种替代，是描述当前现象的数值函数的随机变量。这里，我将运用样本空间进路，但是，能够把所说的内容轻易地转化为一种随机变量的观点。

从一种形式的观点来看，样本空间 Ω 能被看成是由 0 和 1 组成的所有无穷序列的集合，而且，这个无穷序列中只含有一个 1。在每个序列中，这一个 1 出现在这个序列的第 n 项，代表一个粒子在第 n 次试验时或在第 n 个时间周期内发生衰变，其中，每次试验或周期都被理解为与其他试验的持续时间一样。于是，设 E_n 是在试验 n 中衰变的事件。设 W_n 是前 n 次试验中没有衰变的事件，所以，

$$W_n = -\bigcup_{i=1}^{n} E_i$$

208 独特结构的公理体现在下列定义中。这个公理恰好断言,已知衰变还没有发生,那么在 n 次试验中衰变的概率等价于在第一次试验中衰变的概率。从而以一种简单的方式表达了整个期间衰变倾向的恒定性或不变性的定性原理。

定义 3. 设 Ω 是由 0 和 1 组成的所有序列的集合,而且,每个序列只有一个 1,再设 $\mathfrak{I}$ 是在含有柱集代数的 Ω 上的最小的 σ-代数。一个结构 $\Omega=(\Omega, \mathfrak{I}, \succeq)$ 是独立于过去的定性的等待时间结构(waiting-time structure),当且仅当,Ω 是一个定性的条件概率结构,此外,倘若 $W_{n-1} \succ \varnothing$,那么,对于每个 n,都满足下列公理:

等待时间公理:$E_n \mid W_{n-1} \approx E_1$。

由定义 3 描述的这种结构被称为独立于过去的等待时间结构,因为这种描述短语刻画了从像衰变现象之类的特殊应用中抽象出来的它们的一般特性。一种有用的更简短的描述是无记忆的(memoryless)等待时间结构。

这种独特结构的公理的简单性,可能与概率的主观理论包括的公理特征(第 7 节),形成了显明的对比。此外,这种表征定理的自然形式是不同的。强调的是满足这些结构公理——在这个案例中,是强调等待时间公理——而且,强调有一个惟一的参量形式,而不是一个惟一的分布。

定理 2. (衰变的表征定理)。设 $\Omega=(\Omega, \mathfrak{I}, \succeq)$ 是一个与过去无关的定性的等待时间结构。于是,在 $\mathfrak{I}$ 上存在着一个概率测度,使得满足等待时间公理,即

(i) $P(E_n \mid W_{n-1}) = P(E_1)$,当且仅当,$E_n \mid W_{n-1} \approx E_1$ 而且,存在着一个数 $p(0 < p \leqslant 1)$,使得

(ii) $$P(E_n) = p(1-p)^{n-1}$$

此外，满足(i)的任意概率测度都有形式(ii)。

证明： 事件 E_n 惟一地确定一个原子或 Ω 的可能实验结果 ω，即对于每个 n，在 Ω 中都有一个 ω，使得

$$E_n = \{\omega\}$$

一种在由无穷序列组成的样本空间中非常不寻常的情形，因为通常情况下，任何一个 ω 的概率都严格地等于零。

如果 $E_1 \approx \Omega$，那么 $P(E_1) = 1$，这个证明是不重要的。另一方面，如果 $\Omega \succ E_1$，那么，对于每个 n，$(W_n, \cdots, W_1)$是满足标准序列的表征定理假设的一个标准序列。编号是反向的，即 $W_{i+1} \subseteq W_i$ 和 $W_{i+1} \mid W_i \approx W_2 \mid W_1$。[如果$(W_1, \cdots, W_n)$是一个标准序列，那么违反了必要的阿基米德公理的无穷序列$(W_1, \cdots, W_n, \cdots)$也是如此。][1]从 W_i 的定义来看，$W_{i+1} \subseteq W_i$ 是明显的。借助于等待时间公理

209

$$W_{i+1} \mid W_i \approx W_1$$

既然 $E_i = W_{i-1} - W_i$ 和 $W_1 = -E_1$，因此，

$$E_1 \approx E_{i+1} \mid W_i \approx W_i - W_{i+1} \mid W_i$$

$$\approx - W_{i+1} \mid W_i,$$

于是，通过基本的运算，

$$W_1 \approx - E_1 \approx W_{i+1} \mid W_i$$

运用标准序列的表征定理，我们知道，存在着一个数值函数 P' 和数 c 与 q $(0 < q < 1$ 和 $c > 0)$，使得

$$P'(W_i) = cq^i$$

---

① 阿基米德公理是必要，因为这个事件代数不是有穷的。

[设 i' 是对相反的顺序 $(n, \cdots, 1)$ 的编号;那么 $i' = n-(i-1)$,而且,指数 $n+1-i$ 在这个表征定理中成为 i'。]从长度为 n 的一个固定的标准序列出发,能够以一种明显的方式把 P' 推广到每个 $i>n$ 的情形。

下一步是通过下列方程,把作为一个相加集合函数的 P' 推广到所有的原子事件 E_i,

$$\begin{aligned} P'(E_i) &= P'(W_{i-1} - W_i) \\ &= P'(W_{i-1}) - P(W_i) \\ &= c(q^{i-1} - q^i) \\ &= cq^{i-1}(1-q) \end{aligned}$$

这种推广的一致性和惟一性很容易核实。现在

$$\sum_{i=1}^{\infty} P'(E_i) = c$$

因此,我们设置 $c=1$,以便从 P' 获得在 1 上赋范的一个测度 P,并设 $p = 1-q$。我们就有

$$P(E_i) = p(1-p)^{i-1}$$

$$P(W_i) = (1-p)^i$$

刚才定义的函数 P 等价于在 Ω 上的一个离散密度,因为对于 Ω 中的某个 ω,$E_i = \{\omega\}$,因此,P 可以通过众所周知的方法惟一地推广到 Ω 的柱集的 σ-代数。

最后,(ii)的证明是直接的。如果我们假设,存在着一个 n,使得 $P(E_n) \neq p(1-p)^{n-1}$,其中,$p = P(E_1)$,我们就可以用等待时间公理得到一个矛盾,因为根据假设 $P(W_{n-1}) = (1-p)^{n-1}$ 和 $P(E_n \mid W_{n-1}) = p$,由此 $P(E_n) = pP(W_{n-1}) = p(1-p)^{n-1}$,即几何分布的标准密度。　　证毕。

我完全承认，对这个特殊表征定理的批评是，从形式的观点来看，倾向性的分析太接近概率分析，我早在关于倾向性的某个文献中提出了这种抱怨。一般情况下，倾向性不是概率，而是提供了由以构造概率的要素。我认为，我所做的有利的肯定论证，不得不被置于一种更敏感因而更脆弱的基础之上。观点已经提出，但我将再一次强调倾向性是如何进入的。无记忆的等待时间公理，是在作为一个基本公理的主观概率的标准理论中，从来不会遇到的一个结构公理。它是相对于特定物理现象的一个特殊公理。因此，它代表了对倾向性的一种定性表达。第二，我们从表征定理中得到的概率不是惟一的，而只是惟一地确定衰变参量。此外，它不是一个主观概念，而完全是一个客观概念。鉴别和定位所确定的物理参量的个数，是对下列观点的一种强调方式：倾向性已经进入，但一直没有给出含有惟一测度的纯粹的概率理论。 210

离散定性的密度。把定性的公理主要限于离散的原子，能够很好地放大刚才提出的定性的观点，因此而极大地简化了定性的公理。

正如在所有类型的基本概率的事例中常见的那样，当一种分布有一种已知的形式时，通常，通过事件的密度比通过事件的概率测度更加容易描述这种分布。在离散的情形中，这种情况在形式上相当简单。每一个原子，即每一个原子事件，都有一个定性的概率，而且，我们只需要判断这些定性概率之间的关系。我们要求一个离散密度的函数 P 表现出下列三种特性：

(i) $p(a_i) \geqslant 0$

(ii) $\sum_{i=1}^{n} p(a_i) = 1$

(iii) $p(a_i) \geqslant (a_j)$，当且仅当，$a_i \geq a_j$

注意，a_i 不是一个实验中的对象或刺激物，而是定性的原子事件，是穷举的和相互排斥的。

我们也需要条件离散密度。为了达到这个目标，我们假设，潜在的概率空间 Ω 是有穷的或可数的，具有在已知事件族 F 上的概率测度 P。对于 a_i 来说，密度 p 和测度 P 之间的关系是 Ω 的一个原子事件，

$$p(a_i) = P(\{a_i\})$$

于是，如果 A 是任何一个事件，使得 $P(A) > 0$，

$$p(a_i \mid A) = P(\{a_i\} \mid A)$$

当然，现在 $p(a_i | A)$ 本身就是一个离散密度，满足(i)—(iii)。下面是这条进路的某个简单但有用的例子。

均匀密度。在定性排序 $\geq$ 中，在 Ω 上的均匀密度通过等价的所有原子事件加以描述，即

$$a_i \approx a_j$$

211 我们很容易表明，满足这种等价性和(i)—(iii)的惟一的密度是

$$p(a_i) = \frac{1}{n}$$

其中，n 是在 Ω 中的原子事件数。

几何密度。我们能够极大地简化导向定理 2 的论证。空间 Ω 是针对定理 2 的。它是离散的但可数的原子事件(a_1, …, a_n, …)的集合，每个原子事件都有状态发生改变的一个正的定性概率。原子事件的编号直观上对应于实验次数。原子事件通过关系 $\geq$，以定性概率进行排序。我们也引入了一个受限制的

条件概率。如果 $i > j$，那么，$a_i | A_j$ 是条件原子事件：已知在试验 j 或 j 之前不发生状态的改变，试验 i 将发生状态的改变。(注意，像前面一样，这里的 A_j 意味着从试验 1 到试验 j 状态不发生任何改变。)定性概率排序关系也能被推广到包括这些特殊条件的事件。

除了前面已知的(i)—(iii)之外，两个假定的特性是：

(iv) 排序特性：$a_i \succeq a_j$，当且仅当，$j \geqslant i$；

(v) 无记忆特性：$a_{i+1} \mid A_i \approx a_1$。

很容易证明，(iv)意味着$\succeq$的一个弱排序。于是，我们能证明 $p(a_n)$有几何密度的形式

$$p(a_n) = c(1-c)^{n-1} \quad (0 < c < 1)$$

但当然，公理(i)～(v)没有确定参量 c 的值。

可交换性。第 1 节最后已经提到可交换的(即在试验数排列条件下不变的)结果序列。它们在这里也应归于许多过程的一个重要的物理不变量，但进一步的分析推迟到关于主观概率的下一节，在那里，分析了德·菲内蒂提出的关于可交换序列的基本定理。

平稳性。在表征许多基本的物理过程时，随时间平稳的过程是普遍的，比如，平衡态的统计力学的过程。正如在 4.5 节中已经明确的那样，任何一个各态历经的过程必定都要求是平稳的。对于离散的状态集和离散时间指标的过程来说，平稳性有一个简单的定性表述：对于试验 $n_1 < n_2 < \cdots < n_m$ 的任何一个有穷序列和任何一个正整数 k，

$$a_{n_1} a_{n_2} \cdots a_{n_m} \approx a_{n_1+k} a_{n_2+k} \cdots a_{n_m+k}$$

其中，a_{n_1} 是试验 n_1 的离散结果，依此类推。平稳性的本质是在时间转变，比如 k，F 的不变性。

反应的倾向性。在心理学中，提出各种不同的理论模型来
212 表征反应强度，有着悠久的历史，反应强度反过来成为选择概率的基础。我所说的"选择概率"是指，在某个已知实验或自然环境中，从一个集合 A 选择一个已知项 a 的概率。通过一个概率表示这种选择的事实反映出：期望效用的标准代数模型不足以代表现实的行为。无论人们可能认为个体应该做什么，有实验和其他环境中的大量证据表明，当他们面对看来是重复相同的可替代的选择集合时，他们不会重复作出同样的选择，这是一个生活事实。在过去的几十年里，对这些情境的形式研究也是一项高强度的工作。最简单和优雅的模型之一是卢斯提供的选择模型(Luce, 1959)。在卢斯自己的发展中，他从根据可观察概率陈述的他的选择公理前进到反应强度的存在性。为了举例说明我们这里所希望的这种思想，我们将从另一种方式开始，即通过假定一个反应强度，然后表明，我们如何从这个假定轻易地推出他的选择公理。第二，在卢斯最初表述的学习模型中可以看出反应强度比反应概率有的一个优势，在这里，从一个试验到另一个试验表现学习的变换，当根据反应强调阐述时，是线性的，当根据选择概率阐述时，却是非线性的。

我现在转向这些思想的形式发展。在预期的解释中，T 表现为是可替代的选择集合，数值函数 v 是对反应(刺激)强度的测量。

定义 4. 设 T 是一个非空集合，设 v 是在 T 上定义的一个非负的实值函数，使得至少对于 T 中的一个 x，$v(x) > 0$ 和 $\sum_{x\in T} v(x)$ 是有穷的。于是，$\mathrm{T} = (T, v)$ 是一个(选择的)反应强度模型。

为一个反应强度模型增加的一般必要条件显然很弱，但我们已经能在下列意义上证明一个表征定理，它比前面的表征定

理更特殊,除了满足有穷相加概率空间的公理之外,也满足卢斯的选择公理。此外,我们为了得到各种不同的学习模型,强加进一步的条件。

定理 3. (表征定理)设 $\mathcal{T}=(T, v)$ 是一个反应强度模型,而且,相对于 $P(T)$ 中的 U,定义 T 的幂集,

$$P_T(U)=\frac{\sum_{x\in U}v(x)}{\sum_{x\in T}v(x)}$$

于是,$(T, P(T), P_T)$ 是一个有穷相加的概率空间。此外,概率测度 P_T 满足卢斯的选择公理,即对于 P_T 中的 V 来说,由于 $\sum_{x\in V}v(x)\neq 0$ 和 v'是把 v 限于 V 中,$\upsilon=(V, v')$ 是一个反应强度模型,使得对于 $U\subseteq V$,

$$P_V(U)=P_T(U\mid V)$$

证明: 这个定理的一般表征部分是明显的。为了证明卢斯的公理,我们注意到,因为 $U\subseteq V$,所以, 213

$$\begin{aligned}P_T(U\mid V)&=\frac{P_T(U\cap V)}{P_T(V)}\\&=\frac{\sum\limits_{x\in U}v(x)}{\sum\limits_{x\in T}v(x)}\Bigg/\frac{\sum\limits_{x\in V}v(x)}{\sum\limits_{x\in T}v(x)}\\&=\sum_{x\in U}v(x)\Bigg/\sum_{x\in V}v(x)\\&=P_V(U)\end{aligned}$$

证毕

在这个定理中所用的符号稍有点微秒,开始时可能会被混淆。

注意,标有下角标的集合代表所有的选择替代的物理集合。在$P_T(U|V)$的情况下,V上的条件集合是关于在T的一个子集的项中选择出的实际出现的信息。卢斯的公理应该成立,这完全不是同义反复。事实上,关于如何最好地估计所观察的选择概率的反应强度,有相当大量的统计文献(Bradley and Terry, 1952; Bradley, 1954a, b, 1955; Abelson and Bradley, 1954; Ford, 1957)。

为了举例说明根据倾向性表述理论如何能够简化某些分析,我们概述卢斯最初的学习模型的情境。设f是把有穷的反应强度向量$v=(v_1, \cdots, v_r)$从一个试验映射到下一个试验的学习函数。这里,我们假设T是有穷的——特别是有基数r。我们假设,反应强度是**无约束的**,即对于任何一个实数α来说,都有一个n,使得$|f^n(v)|>\alpha$,其中,f^n代表学习函数的n次重复。第二,学习的**叠加**成立,即对于任何一个v, $v^*>0$

$$f(v+v^*) = f(v) + f(v^*)$$

第三,**尺度或单位的独立性**成立,即对于$v>0$和任何一个实数$k>0$,

$$f(kv) = kf(v)$$

但在线性代数中一个众所周知的结果是,假设条件意味着f在已知的r维的向量空间上是一个线性算符。相反,在这些而不是更强的假设条件下,从一个试验到另一个试验表现出的反应概率$P_T(U)$是复杂的;特别是,它们不能通过概率的线性变换关联起来。

尽管定理3的证明很简单,而且,在迄今为止的发展中,几乎没有把结构强加给这个反应强度函数,但是,预期解释与概率的倾向性概念很符合——至少在我面对各种倾向性的表征定理

的发展时是如此。事实上,我喜欢的倾向性的多元论的一面是,根本不存在独特的自然表征定理。许多不同的物理学和心理学的倾向性应该产生独特的表征定理。另一方面,定理 3 和根据某种倾向性的概率的其他类似的"直接"表征的一个明显缺陷是,无法确保提供随机性。这也是放射性衰变事例的不足。在这两种情况下,增加公理研究随机性是困难的。在衰变的情形中,顺其自然的是,必须研究实时现象,而不是典型的多项式的情形。在反应强度模型中,人们期望直接从重复中学习,因而明显的试验序列将不代表平稳过程,原则上,正如在第 4 节所分析的那样,能运用定义有穷随机序列的标准机制,但这里不重复其细节。 214

出现正面的倾向性。至少从彭加勒(1912)的分析物理系统开始,就存在着一种传统:我们通常把赌博装置看成是经典力学系统。斯莫卢乔夫斯基(Smoluchowski, 1918),特别是,霍普夫(Hopf, 1934)提供了赌博装置的更详尽的应用。冯·柏拉图(von Plato, 1983)令人满意地记载了这些思想史。这里提出的只需要黎曼积分的简单进路主要归功于凯勒(Keller,1986)。

我们在几个简化假设的条件下分析掷硬币的机制,但如期所料,在实践中,这些假设只部分地得到满足。第一,我们考虑一个半径为 a 的圆形硬币,硬币的厚度忽略不计。第二,我们假设,硬币的金属材料的分布是非常均匀的,所以,它的引力中心也是它的几何中心。因此,假设忽略不计正反面标记的差异。第三,我们忽略硬币在自旋或下落时产生的任何摩擦。第四,我们只分析落在桌面上的第一个作用点。我们假设,以此为出发点,硬币一直是正面向上。由此我们可忽略硬币在静止下来之前,自旋、反弹和再次落下的任何一个弹性问题。正如霍普夫指出的那样,用我们能够阐述的最现实的假设替代上面的所有假

设的数学分析，仍然不能以完备的形式获得，真正的系统是耗散的，而不是守恒的。另一方面，我们作出的理想化不是完全不切实际的；有好的物理学理由认为，由于摩擦等原因，关于耗散的更现实的假设，完全不影响数学分析的概念特征，而是影响定量细节，就我们的目的而言，定量的细节不是关键。

同样有用的是质问，这种已知的分析是否符合标准的质点力学理论。回答是，"几乎符合，但不完全符合"。以旋转为例，我们把硬币看成是一个刚体，而不是一个质点，尽管我们恰好通过一个质点集合(质点间的相互距离保持不变)能够模拟自旋特性。

现在转向更加形式的细节。我们用笛卡儿坐标系，x 和 z 在水平面上，y 是高度的测量，这样，$y(t)$是硬币在时间 t 的引力中心的高度。只有垂直方向的力是引力，因此，牛顿运动方程是

$$\frac{d^2y(t)}{dt^2}=-g \tag{4}$$

其中，g 是不变的引力加速度。我们假设，在时间 $t=0$，初始条件为，高度是 a，向上投掷硬币的速度是 u，即

$$y(0)=a,\ \dot{y}(0)=u \tag{5}$$

215 方程(4)和(5)惟一地确定 $y(t)$直到作用点。很容易证明，

$$y(t)=-\frac{gt^2}{2}+ut+a \tag{6}$$

至于自旋，我们假设，硬币是围绕一条水平轴旋转，这条水平轴是硬币的一条直径；我们确定 z 坐标轴平行于这条旋转轴，而且，我们假设角度的位置如下。角 $\theta(t)$是正的 y 轴和硬币正面的一条垂线之间的夹角——正如在图 1 中看到的那样，两条线都在 xy 平面，图 1 来自凯勒(1986)。我们假设，最初，这个

硬币是正面水平向上，以正的角速度 ω 投掷。因此，在 $t=0$，初始自旋条件是：

$$\theta(0)=0,\ \dot{\theta}(0)=\omega \quad (7)$$

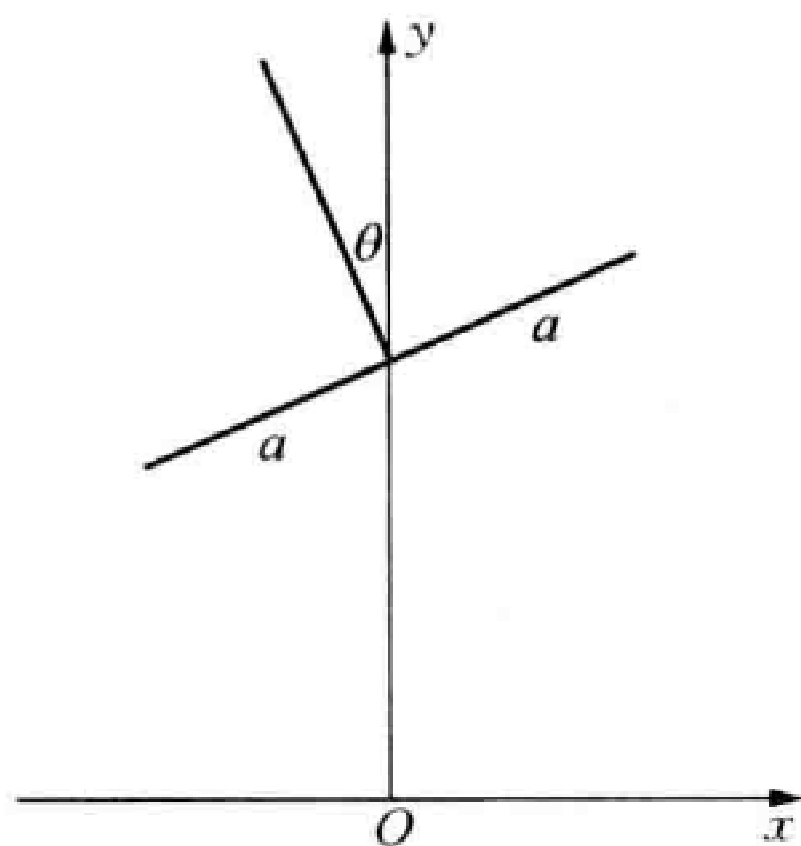


图 1　xy 平面与直径长为 $2a$ 的硬币相交。正常硬币的正面与正 y 轴的夹角为 θ。

此外，假设没有耗散，正如我们所做的那样，支配硬币旋转运动的方程恰好是常速度的方程

$$\frac{d^2\theta(t)}{dt^2}=0 \quad (8)$$

(7)和(8)的惟一解是

$$\theta(t)=\omega t \quad (9)$$

设 t_s 是硬币落下时触及平面时的时间，我们把这个平面看成是 $y=0$。已知前面陈述的假设：硬币根本不会反弹，硬币正面向上，当且仅当，

$$2n\pi-\frac{\pi}{2}<\theta(t_s)<2n\pi+\frac{\pi}{2},\ n=0,\ 1,\ 2,\ \cdots \quad (10)$$

现在我希望找到 t_s。首先，我们注意到，在任何时间 t，硬币的最低点在 $y(t)-a\mid\sin\theta(t)\mid$。因此，$t_s$ 是下列方程的最小的正根　216

$$y(t_s)-a\mid\sin\theta(t_s)\mid=0 \quad (11)$$

我们接下来希望找到凯勒所说的**正面的预映射**(pre-image)，即找到硬币最后正面向上时，速度 u 和角速度 ω 的初始值的集合。假设我们称之为在 $u\omega$ 平面 H 上的点集。

我们首先看到通过(10)定义的终点。这些与(9)一起产生——只是对终点而言——，

$$ut_s = \left(2n \pm \frac{1}{2}\right)\pi \tag{12}$$

而且同样,在这些终点, $\sin\theta(t_s) = \pm 1$, 因此,在这些情况下把(11)简化为

$$y(t_s) - a = 0 \tag{13}$$

它与(8)合并得出,

$$ut_s - \frac{gt_s^2}{2} = 0 \tag{14}$$

终点的重要性是,它们确定了区域 H 的边界。特别是,我们考察(14)的解来确定 t_s,然后,考察(12)的解来确定边界。方程(14)有两个解:

$$t_s = 0,\ t_s = \frac{2u}{g}$$

第一个解只产生平凡的结果,因此,我们只根据 ω 和 u,运用(12)中的第二个解,获得一种关系:

$$\omega = \frac{\left(2n \pm \frac{1}{2}\right)\pi g}{2u},\ n = 0,\ 1,\ 2,\ \cdots \tag{15}$$

对于 n 的不同值,关系(15)图示为图 2(凯勒之后,1986)。正如从(15)能够看到的那样,每条曲线都是一条双曲线。在这个轴上 $\omega=0$,投掷时保持正面向上,这样,条状以属于 H 的轴为边界。下一个条状是 T 的一部分,即 H 的补集,显然,条状一直轮流交替的要么是 T 的一部分,要么是 H 的一部分。从(15)来看,我们能推出,这些条状是等垂直分隔的,即$\frac{\pi g}{2u}$, $n=0$ 除外,因为离这个轴的垂直距离最短的间距是$\frac{\pi g}{4u}$。

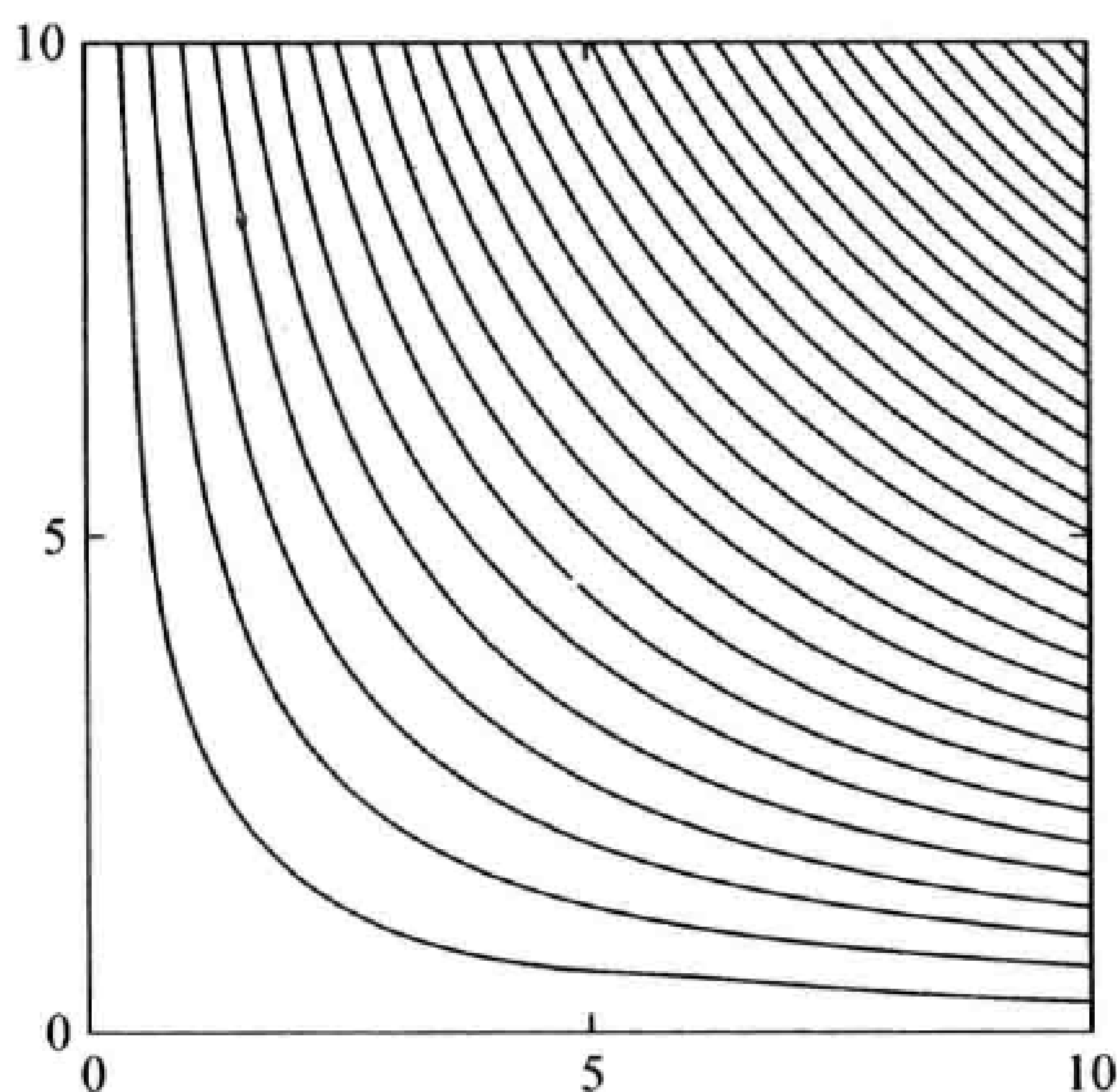


图 2　由于横坐标为$\frac{u}{g}$,所以 n 的不同值显示了把集合 H 和 T 分开的曲线,即在初始条件的 $u\omega$ 平面上的正反面的预映射。

关键的观察是,当加大投掷的初速度时,垂直的间距就减小和趋于零。这意味着,当 u 增加时,H 和 T 之间的交替是通过很小的速度变化产生的。

正如我们将要表明的那样,基本上任何一个数学上可接受的 u 和 ω 在 $t=0$ 的概率分布,都将导致出现正面的概率近似于 0.5。投掷的力学过程支配着结果。在 u 和 ω 中完全不可避免很小的初始变化,都导致标准的可能结果。更明确地说,掷硬币的标准的力学过程,有一种强的倾向性是,以 0.5 的概率产生正面向上的结果。

为了计算 P_H,即出现正面的概率,我们假设,对于 $u>0$ 和 $\omega>0$ 来说,初始的连续概率密度为 $p(u,\omega)>0$。重要的是注意到,对于 $p(u,\omega)$,不能假设任何对称性的特征。我们立即有

217

$$P_H = \iint_H p(u,\ \omega)\mathrm{d}\omega\mathrm{d}u \tag{16}$$

显然,(16)没有对 P_H 强加任何限制。已知所希望 P_H 的任何一个值,即出现正面的任何一个概率,我们就能按照(16)找到产生它的一个条件密度。

所能证明的是,在 $P_H = \frac{1}{2}$ 的极限范围内,增加速度 u。实际的收敛速率对已知密度 $p(u,\ \omega)$ 很敏感。

定理 4. (表征定理)

$$\lim_{U\to\infty} P(H \mid u > U) = \frac{1}{2}$$

证明: 我们首先写出没有采用极限的条件概率

$$P(H \mid u > U) = \frac{\int_U^\infty \sum_{n=0}^\infty \int_{(2n-(1/2))\pi g/2u}^{(2n+(1/2))\pi g/2u} p(u,\ \omega)\mathrm{d}\omega\mathrm{d}u}{\int_U^\infty \int_0^\infty p(u,\ \omega)\mathrm{d}\omega\mathrm{d}u} \tag{17}$$

218 [在分子中所要做的是,对(15)提供的 H 的每个"片段"进行积分。]集合 T 是 H 的补集,因此,T 的片段的边界是

$$\omega = \frac{\left(2n + 1 \pm \frac{1}{2}\right)\pi g}{2u} \tag{18}$$

而且,我们能把(17)的分母写为

$$\int_U^\infty \sum_{n=0}^\infty \int_{(2n-(1/2))\pi g/2u}^{(2n+(1/2))\pi g/2u} p(u,\ \omega)\mathrm{d}\omega\mathrm{d}u$$
$$+ \int_U^\infty \sum_{n=0}^\infty \int_{(2n+(1/2))\pi g/2u}^{(2n+(3/2))\pi g/2u} p(u,\ \omega)\mathrm{d}\omega\mathrm{d}u \tag{19}$$

我们接下来希望估计,当 $U\to\infty$ 时,(17)右边的分子是什么。从黎曼积分的定义来看,我们能够根据间隔的长度,即间隔

是 $\pi g/2u$，在求和的那边，接近每个积分，而且，在这个间隔的中点，即 $p\left(u,\ \frac{n\pi g}{u}\right)$，积分值是

$$
\begin{aligned}
&\lim_{U\to\infty}\int_{U}^{\infty}\ \sum_{n=0}^{\infty}\int_{(2n-(1/2))\pi g/2u}^{(2n+(1/2))\pi g/2u} p(u,\ \omega)\mathrm{d}\omega\mathrm{d}u \\
&=\lim_{U\to\infty}\int_{U}^{\infty}\ \sum_{n=0}^{\infty} p\left(u,\ \frac{n\pi g}{u}\right)\frac{\pi g}{2u}\mathrm{d}u \qquad (20)
\end{aligned}
$$

直接表明，当 $U\to\infty$ 时，右边的积分收敛于 $\frac{1}{2}P(u>U)$。从结果(17)、(19)和(20)来看，我们有所需要的

$$
\begin{aligned}
&\lim_{U\to\infty}P(H\mid u>U) \\
&=\frac{\lim\limits_{U\to\infty}\frac{1}{2}P(u>U)}{\lim\limits_{U\to\infty}\frac{1}{2}P(u>U)+\lim\limits_{U\to\infty}\frac{1}{2}P(u>U)}=\frac{1}{2}
\end{aligned}
$$

如果不追求细节，有价值的是，在通常掷硬币时所实际观察到的范围内，只基于线性角速度，略述一个更现实的近似分析。为了做到这一点，而不是接受一个物理上不现实的极限，我们对密度 $p(u,\ \omega)$ 作出更多的限制：

(i) 密度 $p(u,\ \omega)$ 是单峰的。

(ii) $p(u,\ \omega)$ 和平行于 $(u,\ \omega)$ 坐标的任何一个固定高度的平面的相交都是凸形的。

(iii) 密度 $p(u,\ \omega)$ 只在 $u\omega$ 平面的一个有穷区域上是正的。特别是，存在着一个 u^* 和一个 ω^*，使得如果 $u>u^*$ 或 $\omega>\omega^*$，则 $p(u,\ \omega)=0$。

(iv) 要求 u 和 ω 的变化量有更低的下限，以致没有更多地集中于图 2 的一个小区域，从而产生出太多与反面相关的正面，

或者相反。

由于选择(i)、(iii)和(iv)在经验上是合理的,所以,能够为正面或反面的概率确定一个非常好的边界,即接近于二分之一。提出在 $p(u, \omega)$上附加的一些滤波条件有利于促进计算;甚至一个二元的正态分布也与真正的数据符合得相当好。

三体运动中随机倾向。第四个例子是三体问题的一种特殊情况,在力学史上,当然是一个受到最广泛研究的问题。我们的特殊情况是这样的:有两个质量(m_1 和 m_2)相等的质点,按照牛顿的反平方万有引力定律,在一个相对于它们的共同质心是椭圆的轨道上运动,质心是静止的。第三个质点的质量几乎可忽略不计,因此,它不影响其他两个质点的运动,而这两个质点会影响它的运动。第三个质点沿着垂直于头两个质点运动平面的一条直线运动,这条直线与它们的质心的平面相交——设这是 z 轴。由于考虑到对称性,我们能够看到,第三个质点不会离开这条线。在莫泽(Moser,1973)之后,参见图3。受约束的问题是描述第三个质点的运动。

219

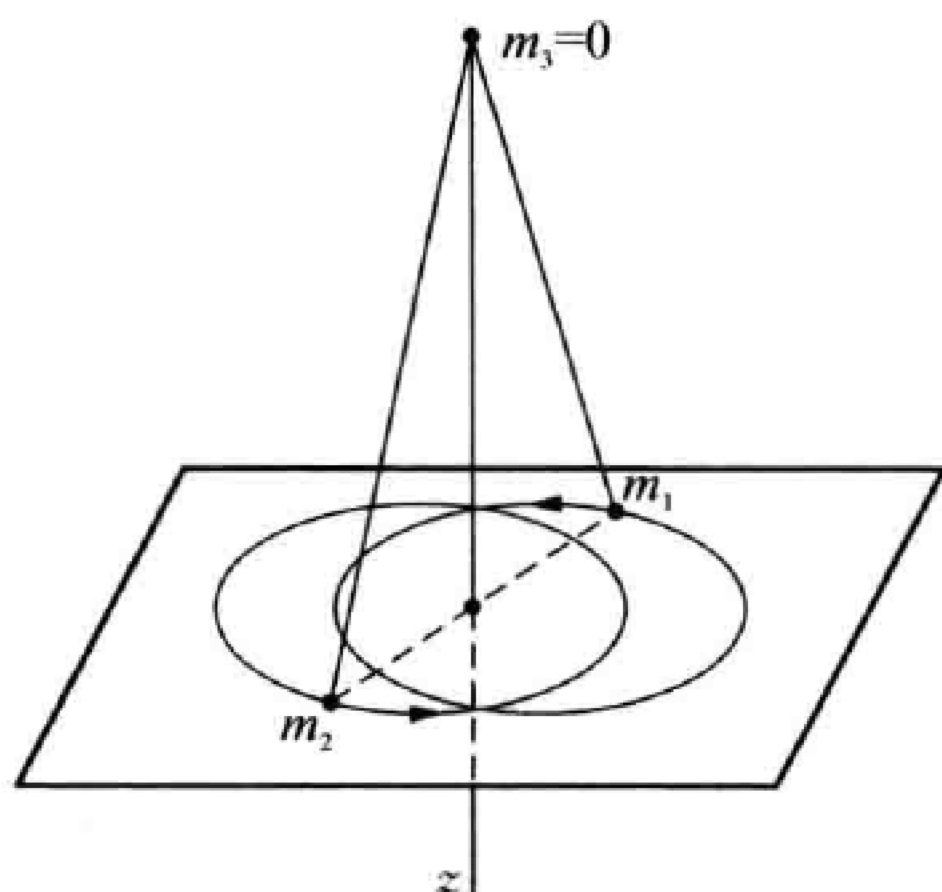


图 3　受约束的三体问题

为了以简单的方式获得微分方程,我们使时间单元标准化,以致两个质量在 xy 平面上旋转的时间周期是 2π,我们采纳使万有引力常数为 1 的长度单元,最后,$m_1 = m_2 = \frac{1}{2}$,所以,$m_1 + m_2 = 1$。根据质点 1 的质量,施加于有趣的质点 3 的力是

$$F_1 = \frac{m_1}{z^2 + r^2} \cdot \frac{(z, r)}{\sqrt{z^2 + r^2}}$$

其中,r 是在粒子 1 的 xy 平面上离两质点系统 m_1 和 m_2 的质心的距离,当然,在 xy 平面上,这个质心恰好是点 $z=0$。注意,$\frac{(z, r)}{\sqrt{z^2+r^2}}$是力 F_1 的方向矢量①单位。同样,

$$F_2 = \frac{m_2}{z^2 + r^2} \cdot \frac{(z, -r)}{\sqrt{z^2 + r^2}}$$

因此,通过简化,我们获得作为第三个质点的一般微分方程:

$$\frac{d^2 z}{dt^2} = -\frac{z}{(z^2 + r^2)^{3/2}}$$

分析这种容易描述的情境是相当复杂的和技术的,但用非形式的术语陈述某些结果,是很简单的。当第三个质点接近逃逸速度时——即质点离开后不再周期性地返回的速度,周期运动是很不规则的。特别是能够证明下面值得注意的定理。设 t_1,t_2,…是第三个质点与其他两个质点运动的平面相交的时间。设 s_k 是等于或小于 t_{k+1} 和 t_k 之差乘以一个常数的最大整数。② s_k 的变化显然测出了这个周期运动的不规则性。正如莫泽(1973)阐述的那样,归功于俄罗斯数学家西特尼科夫(Sitnikov,1960)和阿列克谢耶夫(Alekseev,1969,b)的这个定理如下。 220

定理 5. 已知椭圆轨道的偏心率是正的,但不太大,存在着一个整数,比如说 α,使得项 s_k($s_k \geqslant \alpha$)的任何一个无穷序列,都

---

① "vecotor"这个词在数学中通常译为"向量",在物理学中通常译为"矢量",本文根据语境同时使用了这两种译法。——译者

② 这个常数是两个质点在平面上运动周期的倒数。

对应于支配第三个质点运动的决定论的微分方程的一个解。①立即得出关于随机序列的一个推论。设 s 是正反面的任何一个随机(有穷的或无穷的)序列——为了达到这个目的,我们能够使用几个不同定义——丘奇、柯尔莫哥洛夫、马丁·勒夫,等等——中的任何一个定义。我们选取大于 α 的两个整数代表这个随机序列——比如说,两个中较小的表示正面,另一个表示反面。我们因此有:

推论 1. 正反面的任何一个随机序列都对应于支配第三个质点运动的决定论的微分方程的一个解。

换言之,对每个随机序列来说,都存在着决定这个相应解的一组初始条件。注意,在根本的方面,这个质点的运动完全是不可预言的,即使是决定论的。这同时也是相关的随机序列的一个结果。重要的是注意到,与前面掷硬币的例子有所不同,因为在这个三体问题中,没有出现初始条件分布,因而也没有出现关于初始条件的不确定性。在掷硬币的例子中,没有一条轨迹本身显示出这样的随机行为。

在第四个例子中,可以把倾向性表示为,第三个质点表现出只低于其逃逸速度的很不规则的运动的倾向。用通常的术语说,我们可能希望说,这种倾向是向往不规则的倾向,但当然,我们能够说更向往一些解,也就是说,确定的初始条件导致产生随机行为的倾向,在符号动力学的层面使这点明确化。②

关于倾向性的更多评论。加拉沃蒂(1987)对我的观点的评

---

① 微分方程的一个解和一个整数序列之间的这种对应是符号动力学这个术语的来源。这样一种对应思想最初由伯克霍夫(G. D. Birkhoff)于 20 世纪 30 年代提出。

② 关于基于过去十多年的研究,对太阳系中的混沌运动的一个新近评论表明,行星的轨道在本质上是混沌的,可参见 Murray and Holman (2001)。

论提出了有关倾向性的许多基本问题,这些问题在许多哲学文献中一直很重要。对我来说,她提出的问题太多,以至于没有希望对所有问题都作出详细的回答,但我认为,有些观点是我能明 221 确接受的,这些观点有助于澄清讨论或至少澄清我自己关于倾向性概念的立场。

倾向性或客观概率。加拉沃蒂首先提出的问题之一是,我的说明是否真的不是对客观机遇的一种说明,反而是倾向性的一种说明。我对这个问题的回答当然是否定的。在我们的观点中,恰好要求一个概率的客观理论是与主观相反的客观,而且,在倾向性说明中蕴含的这种因果性说明是根本没有必要的。例如,纯逻辑的概率理论被预期为是客观的,即与主观的相反。基于相对频率的说明,或者基于有穷序列的复杂度概念的说明,也是如此。这肯定是这些人的一种观点,他们是给出概率的客观理论的这些不同观点的拥护者。此外,所考虑的四个展开的例子,对相对频率或刚才提到的概率的复杂度并没有说什么。似乎对我来说,正是倾向性理论提供一种特殊因果性的概率观,这种概率观有资格作为一种特殊的客观概率。

这就留下了一个有待解决的问题:一种倾向性说明是否包括了概率的所有客观说明。像梅勒(1971)那样的哲学家似乎认为如此,但我持有一种多元论的观点。根据复杂度对概率或随机性的说明,总是能够被还原为根据倾向性的说明,这在原则上是毫无道理的。

细节问题。加拉沃蒂谈到了波普尔、梅勒、吉雷等人关于倾向性的许多详细的哲学讨论。她正确地注意到,我的分析没有探讨这些不同作者所持的不同观点的细节。我试图在不同的意义上,即在呈现倾向性的特殊例子的数学细节和科学细节的意义上探究细节。我选择这些例子的一个目的是,举例说明倾向

性概念的多元论。似乎对我来说,加拉沃蒂要求的倾向性概念比我准备要提供的倾向性概念更特殊和更单一。我怀疑,没有对倾向性的一般的技术说明,正如没有对意向(disposition)概念或趋向(tendency)概念的一般的技术说明。在我的观点中,**倾向性**这个术语本身通常不会出现在进展顺利的任何一个科学理论中。它可能出现在对实验的非形式的描述中,但根本不会出现在理论的陈述中。这意味着这个概念不同于在科学理论中确实出现的其他一般概念,但这绝没有对这个概念的有用性提出挑战。关于因果性的概念也能这么说。"**因果性**"这个术语绝对不会或几乎绝对不会出现在详尽的科学理论中,但它肯定出现在围绕这个理论的非形式的交谈中,也通常出现在对实验或观察程序的非形式的描述中。**因果性**这个术语不出现在绝大多数科学理论的表述中,这个事实并不支持这样的论据:因果性概念在科学中是不重要的。恰恰是在理论的表述中,我们研究更特殊的特性和概念。倾向性的情况也是如此,我把这当作是谈论意向的因果性现象。正如许多不同类型的过程或事件是有原因的一样,许多不同类型的过程或事件也是有倾向性的。

222

为了强调这一点,让我只回顾一下前面考虑的四个例子(即衰变的倾向、反应的倾向、掷出正面的倾向和随机的倾向)中的每一个例子,来阐述谈论倾向性的自然方式之一的变化。注意,在头两个例子中,倾向性是指与一个事件或一种行动相关,在后面的两个例子中,是指与一种持续的现象相关,但是,这些文法上的不同与阐述特殊理论的不同方式不一样重要。在这个方面,第三和第四个例子特别令人感兴趣,因为它们都是在经典力学的框架内以一般的方式建立起来的。另一方面,第三个例子要求,根据垂直速度和旋转率的初始条件,引入一种初始概率分布。如果没有体现初始条件变化的这种分布,就肯定不可能完

全从物理学的考虑推出这些结果。第四个例子具有相当不同的特征。没有引入关于概率的附加假设。显示任何一种目标概率的随机序列,完全是通过受约束的三体问题的不同的客观初始条件产生的。我先表明,这如何拒绝了讨论倾向性和非决定论之间关系的许多观点,然后,再回到第四个例子。现在的要点是,第三和第四个例子举例说明了,即使在经典力学的框架内,我们也可能对与概率现象相关的倾向性有非常不同的谈论。这样的多元论的存在也体现在头两个例子的其他方面,使得我怀疑任何一个高深的倾向性都有一个统一的理论,正像我认为我们不可能有关于一般因果性概念的完全令人满意的统一理论一样。例如,我们可能会有一个好的概率因果性的一般理论,但不可能有一个详尽的因果性理论,那就是精确反映不同科学部分的因果现象的相关细微差别的理论。

单一情形与长期频率。不同的哲学家对倾向性的应用各不相同,并对下列问题一直持有相当不同的观点:概率的倾向性进路是应该应用于单一情形的现象,还是只应用于长期频率。我的一般观点是,这是无关紧要的题外话,因而我将不再对此加以讨论。最重要的观点是,该问题主要是一个统计推断,而不是上面提到的一个概率。我乐意用倾向性理论作出单一情形的预言,也作出长期频率的预言,但如何检验这些预言,是统计方法论的问题,而不是概率基础的问题。这是省略这个论题的一种概括方式,我也承认,这不是完全令人满意的,但我希望避免介入到所包括的方法论论题的讨论当中,似乎对我来说,这点对概率的倾向性观点并不特殊。

倾向性和非决定论。从哲学文献中所讨论的观点来看,加拉沃蒂这样说一定是合适的,即倾向性通常蕴含着对非决定论的一种承诺。梅勒(1971)在他的决定论与自然定律这一章的开 223

头指出,“如果倾向性不断地出现,那么决定论就是错误的。”汉弗莱斯(Humphreys,1985)持有同样的观点,他在一篇文章的开头指出,“尽管机遇无可否认地是一种神秘的东西,但是,接近它的一种有希望的方式是通过运用倾向性——在一种特殊环境中,系统拥有的非决定论的意向……”(我在后面返回到由汉弗莱斯的文章提出的其他论题。)我增加第四个例子的理由之一是,为通常分类是完全确定的一个系统,也会出现随机现象提供一个特别详细而高深的实例。注意,在这个例子略述的受约束的三体问题中,根本没有碰撞问题,碰撞有时被用作在经典力学中引入非决定论的基础。在通常意义上,正如在微分方程理论中反映出的那样,在对决定论等问题的哲学讨论中,所描述的三体问题在特征上完全是决定论的,然而,通过适当的初始条件,就能产生任何一种想要的随机性。有时在倾向性和决定论之间的隔离是幻想的。关于非决定论和在决定论框架内密切相关的不稳定性概念,有更多的论述,但在这里展开细节将会使我们走得太远。

倾向性不是概率。在关于倾向性的一篇较好的文章中,汉弗莱斯(1985)对倾向性为什么不可能是概率提供了很好的论证。汉弗莱斯提出的观点是,正如他指出的那样,尽管“一般看法”是倾向性有概率特性,但他认为,情况并非如此,我也同意这一点。他继续把这种论证看成是“拒绝把当前的概率理论看成是正确的机遇理论的一种理由”。我完全不同意他的结论,因此,我要清理一下我赞成和反对他的文章中的不同论点的根据,加拉沃蒂在她的评论中参考了汉弗莱斯的这篇文章。

首先,他根据贝叶斯的定理和逆概率,给出一个形式论据来表明倾向性不可能有这些特性。我将不详细考察这个论据,因为我同意这一点,但在我自己的例子中,我从来没有认为,倾向

性确实有概率的特性。这是我最初拒绝这种观点的理由，也是我拒绝汉弗莱斯的结论的理由。

本章举例说明了，核心观点是建立在表征定理的本性之基础上的。在证明一个表征定理时，根据对象建构一个表征，这些对象本身没有被表征对象的特性，这是正常的。本节广泛地利用的来自经典力学的第三和第四个例子，举例说明了这一点。退一步讲，在哲学界，断言经典力学例示了概率特性，是罕见的。我们在掷硬币的例子中所得到的是，一个任意的初始概率分布和在掷硬币的力学意义上的强对称性的一种结合。从这些不同的因素中，我们获得了实际投掷过程的一种对称分布。运用两个简单的力学微分方程的物理分析，远离概率的任何一个特征。在受约束的三体问题的第四个例子中，这种严格的决定性理论本身，肯定不能被认为是概率理论的组成部分。在力学中最令 224 人吃惊的定理之一是，我们能够在这样一个纯决定论的框架内，以这样一种明确的决定论的方式描写任何一个随机序列。另一方面，在受约束的三体问题中显示出的著名的万有引力的力学性质，完全没有概率结构的性质。

正如我前面所说的那样，第一个例子接近于概率，因而不是下列问题的一个好事例：即从本身不是概率的因素，如何建立根据倾向性的概率表征。第二个例子更接近于概率，因为作为原则上无边界的直观的反应强度概念，并不是一个概率特征。同样的例子是，在和总是不变的一个有穷测量中缺乏恒定性，也不是概率理论的一个特征。

表征定理的其他经典事例表明了同样的问题。其中，最有名的例子是，通过粒子的统计系综表征热力学特性。假如我们说，气体的宏观热力学特性是通过粒子系综描述的，据此，所研究的特性也是不明确的，比如，对一个粒子来说，根本没有富有

意义的温度概念。

因此,正是根据把被表征的对象与建构表征的那些成分分离开来,我拒绝接受汉弗莱斯的这个结论:当前的概率理论应该加以拒绝。正是基于此,我论证了量子力学中的经典概率的概念,但这里考虑量子力学的事例,是不适当的。无论如何,在汉弗莱斯拒绝把标准的概率理论接受为一种适当的机遇理论时,我认为,他没有想到,量子力学独特的概率理论的非标准特征。

沿着相关的思路,在我赞同的萨蒙(Salmon,1979)关于倾向性的评论性文章中也有许多讨论。他似乎一定同意汉弗莱斯的观点:倾向性不是概率。另一方面,他没有像我一直试图做的那样,继续提出倾向性的一种形式的表征理论,而且,在我自己的判断中,为了使倾向性理论达到概率的相对频率理论或主观理论特有的同等水平的技术细节,这是缺少的最重要的应有成分。我也应该说,梅勒著作(1971)的非常缺少的成分是,关于表征定理没有一个明确观点,他的著作不仅对倾向性的解释有许多有用的和令人感兴趣的东西,而且,在具有悠久剑桥传统的关于概率的著作中,缺乏一种表征理论观将带来的一种形式的明晰性。

重新强调多元论。最后,我想对加拉沃蒂评论的最后一段作出说明。她把我举的例子与她同科斯坦蒂尼(Costantini)和罗莎(Rosa)关于基本粒子统计学的工作作比较。她对他们关于基本粒子统计学的现象学特征的分析发表的评论,与对倾向性理论的评论相反,倾向性理论大概试图对各种类型的统计,例如,泊松—爱因斯坦统计和费米—狄拉克统计,如何产生和为什么会产生,给出一种说明性的解释。不幸的是,似乎对我来说,她在这里错过了表征定理的观点和我在倾向性说明中的多元论

的主张。就像对世界的物理特性没有一个深层次的一般说明一 225
样,对倾向性也没有一个深层次的一般说明。在物理学中,我们有许多不同的子学科,每一门学科都对一类现象的特殊性进行深入研究。我们自然不会以任何一种容易的方式跨越整个范围。倾向性的情况也是如此。如果我们为已知现象证明一个表征定理,这将是从对那些现象的一种特殊特征的角度来说的,而且,这种特殊说明,当被转移到另一种环境时将是无效的。在我举的四个例子中已经能看到了这一点。从反应强度到掷硬币,甚至进一步到受限制的三体问题,有很长路要走,可是,在每个例子中,每一种倾向性都形成了能够从概率角度建构一个表征定理的成分。

我主张,根本不存在关于倾向性的准确的一般说明。存在着的是,针对不同的经验情境,多元地揭示关于倾向性的许多说明。提供对倾向性的大概详细的说明,和提供对特性的大概详细的说明一样不可行。

5.7 主观概率的理论

在前面几节,我主要考察了可以说是概率形式理论的"强推广"。在每一个这些案例中,至少有在样本空间上或在已知的随机变量的分布中惟一地确定概率测度的一种建议与方法。我们可以用第二章的一个术语,以稍微更技术的方式,来描述这些强推广企图达到的目的。回想一下,如果一个理论的任何两个模型是同构的,这个理论就是范畴的。此外,如果一个理论的模型有一个相当简单的结构,即一个构成模型域的基本集合,以及在这个域上定义的运算与关系,那么,如果它们的域有相同基数的任何两个模型,是同构的,这个理论被说成是半范畴的。把这些

范畴概念推广到当前的情境是为概率空间定义$(\Omega, \mathfrak{I})$-范畴。我们于是说，拉普拉斯的经典定义、相对频率理论、确证理论和倾向性理论提供了四种不同的概率理论，这些概率理论都是$(\Omega, \mathfrak{I})$-范畴的。我的意思是说，一旦确定了这个样本空间的结构和确定了事件代数，那么这四个理论中的任何一个理论都惟一地(或至少相当于一个小的参数集)确定了这个样本空间上的概率测度。我们也略为详细地考察了这些$(\Omega, \mathfrak{I})$-范畴理论的每一个理论的困难。所有这些理论的一个通病显然是，它们试图做太多的事情。我们希望应用概率概念的许多异类的和不同的情形，使得任何一个$(\Omega, \mathfrak{I})$-范畴理论都是可能的，这一观点似乎是可疑的。

因为我们希望用概率概念来谈论从扑克牌中抽出四个点的机会到明天有雨的概率或在量子力学实验中位置的概率分布的每一个问题，所以，不可能找到一个概率的简单的范畴理论，就不足为奇了。概率的主观理论承认这种应用的多样性，并且事226 实上，用此来论证，为了获得一种概率分布必须处理信息的许多方式，不允许无条件的编纂。结果，在近似相同的环境中，两个有理性的人可能会对还没有观察到的一个事件的概率持有不同的信念。例如，按照概率的主观理论，两位气象学家可能面对相同的气象云图和相同的基本气象变量(比如，表面温度、气压、湿度、高空气压、风，等等)的观察史，可是，他们在预报明天早晨是否下雨时，赋予的数值概率仍然不同。然而，我赶紧补充一句，“主观的”这个术语可能会引起误解。在对主观概率赋值时，支持每一种可能的多元性，并不是概率的主观理论的一个组成部分。主观理论的一个固有而重要的组成部分是分析，例如，如何把经典的相对频率数据合并到主观概率的正确估算中。就像我们能够从概率起重要作用的科学文献的考察中看到的那样，在

对概率的任何一个系统的估算中，这样的数据显然是重要的。在经典定义中自然地应用的对称性原理，在随时可被应用的主观概率的估算中起着重要的作用，这也是显然的。

第1节对贝叶斯定理的讨论，已经提供的一个事例是，把这类强约束置于任何一个主观理论。在不违背主观理论的前提下，不同的研究者选择出的先验概率分布可能在很大程度上是不同的；但如果这些研究者认同获得进一步证据的方法，如果他们可获得共同的观察，那么这些公认的观察通常将使他们的信念趋于一致。

德·菲内蒂的定性公理。让我们现在转向更系统地讨论主观理论的主要问题。关于许多问题的更详细的讨论，读者可参考德·菲内蒂的历史上重要的文章(1937/1964)，在凯伯格和斯莫克勒(Smokler)主编的论文集中已经翻译了这篇文章，也参考德·菲内蒂的文集(1974,1975)。德·菲内蒂的1937年的论文是20世纪在概率的主观基础上最重要的工作之一。自1950年以来，关于这些问题的最有影响的工作大概是萨维奇的著作(1954)。萨维奇通过更多地关注决策行为问题，推广了德·菲内蒂的思想，我将在本节的后面更多地谈到概率的主观理论和决策理论之间的关系。更多哲学的而更少数学的系统著作是由杰弗里(Jeffrey,1965)和利瓦伊(Levi,1980)完成的。

或许，开始对这种主观理论进行系统分析的最佳方法是，考虑德·菲内蒂的定性概率公理。这些公理的精神是，限制概率的定性判断，这种判断足以证明一个标准的表征定理，即确保在标准意义上存在着数值概率测度。从这种观点来看，可以认为，这些公理对测量理论作出了贡献，特别是以概率的比较判断为例。这样一个公理集的中心问题是，为了获得关于事件的一个数值概率测度，*更有可能*的定性关系必须满足多么复杂的条件。

227 使用比较的定性关系的直观思想是，实际上能够指望人们以一种直接的方式作出这样的判断，因为当要求这种比较是定量的时，他们无法作出这样的判断。在大多数情况下，我能明确地说，我认为，在接下来的四个小时内，斯坦福是否更有可能下雨，但我不会以同样直接的方式判断不下雨比下雨的可能性大多少。概括这个例子，接着自然禁不住要问，为了用一个标准的概率测度进行表征，一种定性的比较关系必须有什么样的形式特性。（后面，我评论了人们的定性判断是否确实拥有这些所要求的特性的一些经验文献。）

我们以定性概率结构的概念为出发点，即以在形式上与3.3节的推广结构的那些公理相似的公理为出发点。这个理论的集合论实现是结构$(\Omega, \Im, \succeq)$，其中，Ω是一个非空集，$\Im$是Ω的一个子集族，关系$\succeq$是在$\Im$上的一个二元关系。我一开始就追随了卢斯和苏佩斯(1965)提供的讨论。

定义 1. 一个结构$\boldsymbol{\Omega} = (\Omega, \Im, \succeq)$是一个定性的概率结构，当且仅当，对于$\Im$中的所有$A$，$B$和$C$，都满足下列公理：

S1. $\Im$是Ω上的一个集合代数；

S2. 如果$A \succeq B$和$B \succeq C$，那么$A \succeq C$；

S3. $A \succeq B$或$B \succeq A$；

S4. 如果$A \cap C = \varnothing$和$B \cap C = \varnothing$，那么$A \succeq B$，当且仅当，$A \cup C \succeq B \cup C$；

S5. $A \succeq \varnothing$；

S6. 非$\varnothing \succeq \Omega$。

$\Im$上的第一个公理与有穷相加的概率空间的第一个公理一样。公理S2和S3刚好断言，$\succeq$是$\Im$中的事件的弱排序。公理S4以定性的术语阐述了互斥事件的重要而基本的加法原理。公理S5说，任何一个事件（在弱的意义上）都比不可能的事件更有可

能，而公理 S6 说，确定的事件在严格意义上比不可能事件更可能。以习惯的方式，把严格关系$\succ$定义为：

$$A \succ B,\text{当且仅当,非 } B \succeq A$$

我们可以把最后一个公理陈述为：$\Omega \succ \varnothing$。

为了在某种程度上为这些公理所强加的结构提供更深刻的意义，我们陈述这些公理的一些直观上可取的和所预期的结果。在某些定理的陈述中，方便的是，运用(在较弱意义上)不太可能的关系，这种关系以通常的方式定义为：

$$A \preceq B,\text{当且仅当},B \succeq A$$

第一个定理说，$\preceq$是子集关系的一种推广。

定理 1. 如果 $A \subseteq B$，那么 $A \preceq B$。

证明： 相反，假设，非 $A \preceq B$，即，$A \succ B$。根据假设 $A \subseteq B$，因此存在着一个与 A 不相交的集合 C，使得 $A \cup C = B$，那么，因为 $A \cup \varnothing = A$，我们立即有

$$A \cup \varnothing = A \succ B = A \cup C$$

因而，与公理 S4 对照，$\varnothing \succ C$ 与公理 S5 矛盾。　　证毕。

一些其他的基本属性如下：　　228

定理 2.

(i) 如果 $\varnothing \prec A$ 和 $A \cap B = \varnothing$，那么 $B \prec A \cup B$；

(ii) 如果 $A \succeq B$，那么 $-B \succeq -A$；

(iii) 如果 $A \succeq B$ 和 $C \succeq D$ 和 $A \cap C = \varnothing$，那么 $A \cup C \succeq B \cup D$。

(iv) 如果 $A \cup B \succeq C \cup D$ 和 $C \cap D = \varnothing$，那么 $A \succeq C$ 或 $B \succeq D$。

(v) 如果 $B \succeq -B$ 和 $-C \succeq C$，那么 $B \succeq C$。

因为证明一个定性的概率结构有许多预期的特性是很容易

的,所以,正如在上述定理中所反映的那样,很自然要问到的一个更深刻的问题是,它是否拥有所有的特性,必然保证存在着一个严格一致的数值概率测度 P,使得对于 $\mathfrak{I}$ 中的所有事件 A 和 B 来说,

$$P(A) \geqslant P(B)\text{,当且仅当,}A \geqslant B \tag{1}$$

如果 Ω 是一个无穷集,在适当的意义上很容易表明,定义 1 的公理还没有强到足以保证存在着这样一个概率测度的程度。从任意基数的无穷模型的角度来看,来自模型的逻辑理论的一般证论就足够了;萨维奇给出一个反例(1954,第 41 页)。德·菲内蒂(1951)强调了在有穷情况下获得一种回答的愿望。克拉夫特(Kraft)、普拉特(Pratt)和赛登贝格(Seidenberg)表明(1959),当 Ω 是有穷的时,这种回答也是否定的;事实上,他们为有五个元素因而有 32 个子集的集合 Ω 找到了一个反例。他们的反例的要点是下列的构造。设 $\Omega = \{a, b, c, d, e\}$,再设 ϕ 是一种测度(不是一种概率测度),使得

$$\phi(a) = 4 - \epsilon$$

$$\phi(b) = 1 - \epsilon$$

$$\phi(c) = 2$$

$$\phi(d) = 3 - \epsilon$$

$$\phi(e) = 6$$

和

$$0 < \epsilon < \frac{1}{3}$$

现在,按照这种测度——存在的排序,排列 Ω 的 32 个子集,当然,满足定义 1 的排序。于是,在这种排序中,我们有下列严格

的不等式：

$\{a\} \succ \{b, d\}$，因为 $\phi(a) = 4 - \epsilon > 4 - 2\epsilon = \phi(b) + \phi(d)$；

$\{c, d\} \succ \{a, b\}$，因为 $\phi(c) + \phi(d) = 5 - \epsilon > 5 - 2\epsilon = \phi(a) + \phi(b)$；

$\{b, e\} \succ \{a, d\}$，因为 $\phi(b) + \phi(e) = 7 - \epsilon > 7 - 2\epsilon = \phi(a) + \phi(d)$。

我们也立刻看到，维持这三个不等式的任何一个概率测度 P 都意味着 229

$$\{c, e\} \succ \{a, b, d\}$$

正如刚才通过增加这三个不等式可以看到的那样，在 ϕ 的情况下，

$$\phi(c) + \phi(e) = 8 > 8 - 3\epsilon = \phi(a) + \phi(b) + \phi(d)$$

然而，与 $\{c, e\}$ 和 $\{a, b, d\}$ 不同的集合 A 根据没有下列特性：

$$\phi(\{c, e\}) \geqslant \phi(A) \geqslant \phi(\{a, b, d\})$$

因此，如果不改变其他任何一个不等式，我们能把由 ϕ 产生的排序调整到下列设置范围：

$$\{a, b, d\} \succ \{c, e\} \tag{2}$$

但是，没有一个概率测度能保留(2)，也保留前面的三个不等式，因此，所修改的排序满足定义 1，但不可能通过一个概率测度来表征。

当然，明显的是，通过把特殊的结构假设增加到定义 1 的公理，比如，3.3 节定义 1 的公理 5，有可能保证存在着满足(1)的一个严格一致的概率测度。例如，在有穷的情况下，我们可能要求所有这些原子事件都是等概率的，尽管这诚然是强加了一个

很强的必要条件。

幸运的是,斯科特(1964)已经找到了有穷情形的一个简单的一般解决方案。(克拉夫特、普拉特和赛登贝格已经阐述了,在有穷情况下,存在严格一致的概率测度的充分必要条件,但是,他们的乘法条件是很难理解的。斯科特的处理代表了在明晰性和简单性方面的一个真正收获。)斯科特阐述的中心思想是,把一个代数条件强加给事件的指标函数(或特征函数)。回想一下,一个集合的指标函数恰好是这样的函数:把 1 赋值于这个集合的元素,把 0 赋值于这个集合以外的所有元素。为了简化符号,如果 A 是一个集合,那么我们用 $\mathbf{A}^i$ 指它的指标函数,$\mathbf{A}^i$ 是一个随机变量。这样,如果 A 是一个事件,

$$\mathbf{A}^i(\omega) = \begin{cases} 1 & \text{如果 } \omega \in A \\ 0 & \text{其他情况} \end{cases}$$

斯科特的条件体现在下列定理中,我不提供这个定理的证明。

定理 3. (斯科特的表征定理)设 Ω 是一个有穷集合,$\succeq$ 是在 Ω 的子集上的一个二元关系。在 Ω 上存在一个严格一致的概率测度 P 的充分必要条件如下:对于 Ω 的所有子集 A 和 B,

(1) $A \succeq B$ 或 $B \succeq A$;

(2) $A \succeq \varnothing$;

(3) $\Omega \succ \varnothing$;

(4) 对于 Ω 的所有子集 $A_0, \cdots, A_n, B_0, \cdots, B_n$,如果 $A_i \succeq B_i$, $0 \leqslant i < n$,并且,对 Ω 中的所有 ω,

$$\mathbf{A}_0^i(\omega) + \cdots + \mathbf{A}_n^i(\omega) = \mathbf{B}_0^i(\omega) + \cdots + \mathbf{B}_n^i(\omega)$$

那么 $A_n \preceq B_n$。

230 为了举例说明斯科特条件(4)的意义,我们可以看一下它如何蕴

含了传递性。首先,根据假设,$A \succeq B$ 和 $B \succeq C$,任何三个指标函数的必要条件是

$$\boldsymbol{A}^i + \boldsymbol{B}^i + \boldsymbol{C}^i = \boldsymbol{B}^i + \boldsymbol{C}^i + \boldsymbol{A}^i$$

即对于所有的元素 ω,

$$\boldsymbol{A}^i(\omega) + \boldsymbol{B}^i(\omega) + \boldsymbol{C}^i(\omega) = \boldsymbol{B}^i(\omega) + \boldsymbol{C}^i(\omega) + \boldsymbol{A}^i(\omega)$$

因而根据条件(4),

$$C \preceq A$$

因此,通过定义,像所渴望的那样,$A \succeq C$。条件(4)的代数方程恰好要求,Ω 的任何一个元素,即任何一个原子事件,恰好是相同数目的 A_i 和 B_i($0 \leqslant i \leqslant n$)。显然,这个代数条件不可能用定义 1 的简单的集合语言来阐述,因此,代表了一个相当强的条件。

一般的定性公理。在 Ω 是无穷的情况下,已经表明,许多强结构条件是充分的,但不是必要的。例如,德·菲内蒂(1937/1964)和库普曼(Koopman, 1940a, b,1941)分别为了达到存在着 Ω 的划分效果,把一个公理运用于概率上等价的任意多的事件。这个公理与定义 1 的那些公理一起足以证明,数值概率测度的存在。萨维奇(1954)讨论了相关的存在条件。对这些不同条件的一个详细的评论可以在克兰茨等人的著作(1971)的第 5 章和第 9 章找到。

然而,正如苏佩斯和扎诺蒂(Zanotti)所表明的那样,稍微超出指标函数的范围,就能给出在有穷情况和无穷情况下的简单的充分必要条件。这种转移是从一个事件代数到相对于 $\mathfrak{I}$ 的扩展的指标函数的代数 $\mathfrak{I}^*$。代数 $\mathfrak{I}^*$ 恰好是(在函数条件下)包含了 $\mathfrak{I}$ 中的所有事件的指标函数的最小的半群。换言之,$\mathfrak{I}^*$ 是具有下列特性的所有集合的交集:如果 A 在 $\mathfrak{I}$ 中,那

么 $\boldsymbol{A}^i$ 在 $\mathfrak{I}^*$ 中，而且，如果 $\boldsymbol{A}^*$ 和 $\boldsymbol{B}^*$ 在 $\mathfrak{I}^*$ 中，那么 $\boldsymbol{A}^*+\boldsymbol{B}^*$ 在 $\mathfrak{I}^*$ 中。很容易表明，在 $\mathfrak{I}^*$ 中的任何一个函数都是在 Ω 上定义的一个整数值的函数。正是从指标函数推广到整数值函数，证明了称为扩展的指标函数 $\mathfrak{I}^*$ 的元素是合理的，这个函数像指标函数一样是随机变量。

必须把定性概率排序从 $\mathfrak{I}$ 推广到 $\mathfrak{I}^*$，并且，必须考虑这种推广的直观理由。设 $\boldsymbol{A}^*$ 和 $\boldsymbol{B}^*$ 是 $\mathfrak{I}^*$ 中的两个扩展指标函数。于是，有 $\boldsymbol{A}^* \succeq \boldsymbol{B}^*$，就是有，$\boldsymbol{A}^*$ 的期望值等于或大于 $\boldsymbol{B}^*$ 的期望值。现在，这个定性的比较不是关于事件的可能发生的比较，而是关于某些受限制的随机变量的期望值的比较。指标函数本身当然形成了还要更受限制的一类随机变量，但是，它们的期望值的定性比较，在概念上，相当于是事件的可能发生的定性比较。

关于扩展的指标函数的期望值的定性比较，不止有一种思231 考方式，因此，考虑几个例子是有用的。

(i) 假设史密斯正在考虑周末假期飞往两个地方。设 $A_j(j=1,\ 2)$ 是在地方 j 为晴天的事件，B_j 是在地方 j 为暖和天气的事件。史密斯感兴趣的定性比较是，$\boldsymbol{A}_1^i+\boldsymbol{B}_1^i$ 的期望值与 $\boldsymbol{A}_2^i+\boldsymbol{B}_2^i$ 的期望值。很自然会坚持认为，通过 $\boldsymbol{A}_j^i+\boldsymbol{B}_j^i$ 之和，过分地简化了这两个结果的效用。正当的反映是，把两个函数的期望值比作是一个信念问题，不是价值或效用的问题。因此，不管人们对晴天与暖和天气的相对愿望有何感受，打赌说，$\boldsymbol{A}_1^i+\boldsymbol{B}_1^i$ 的期望值将大于 $\boldsymbol{A}_2^i+\boldsymbol{B}_2^i$ 的期望值，是相当自然的。换句话说，在决策理论的语境中，用扩展的指标函数来构造主观概率测度，而不是衡量效用。

(ii) 考虑由 n 个人组成一个特殊人群，从 $1,\ \cdots,\ n$ 对每个人进行编号。设 A_j 是个体 j 在今年到达夏威夷度假的事件，设

B_j 是个体 j 到达阿卡普尔科的事件。于是,定义

$$\boldsymbol{A}^* = \sum_{j=1}^{n} \boldsymbol{A}_j^i \text{ 和 } \boldsymbol{B}^* = \sum_{j=1}^{n} \boldsymbol{B}_j^i$$

显然,$\boldsymbol{A}^*$ 和 $\boldsymbol{B}^*$ 是扩展的指标函数——我们已经离开了蕴含的基础集合 Ω。定性地比较 $\boldsymbol{A}^*$ 和 $\boldsymbol{B}^*$ 的期望值,是有意义的和相当自然的。可以推测,这样的比较事实上对于旅行社、航空运输等来说,是有明确意义的。

我们相信,期望值的这种定性比较,在许多其他语境中,也是自然的。下列的表征定理所表明的是,不管可能结果的集合 Ω 是有穷的,还是无穷的,关于扩展的指标函数的定性比较的很简单的充分必要条件保证了存在着一个严格的一致的有穷相加测量,即 $P(A) \geqslant P(B)$,当且仅当,$A \succeq B$。

这些公理体现在扩展的指标函数的定性代数的定义中。有必要注意几个符号问题。首先,$\boldsymbol{\Omega}^i$ 和 $\varnothing^i$ 分别是可能结果的集合 Ω 和空集合 $\varnothing$ 的指标函数或特征函数。第二,对于在 $\mathfrak{I}^*$ 中的一个函数来说,符号 $n\boldsymbol{A}^*$ 恰好是 $\boldsymbol{A}^*$ 与自身相加 n 次的(函数的)和的标准符号。第三,在 $\mathfrak{I}$ 和 $\mathfrak{I}^*$ 上的排序关系用了相同的符号,因为在 $\mathfrak{I}^*$ 上的排序是在 $\mathfrak{I}$ 上的排序的一种扩展:对于 $\mathfrak{I}$ 中的 A 和 B 来说,

$$\boldsymbol{A}^i \succeq \boldsymbol{B}^i \text{,当且仅当,} A \succeq B$$

最后,以通常的方式把严格排序关系$\succ$定义为:$\boldsymbol{A}^* \succ \boldsymbol{B}^*$,当且仅当,$\boldsymbol{A}^* \succeq \boldsymbol{B}^*$,和非 $\boldsymbol{B}^* \succeq \boldsymbol{A}^*$。

定义 2. 设 Ω 是一个非空集合,设 $\mathfrak{I}$ 是 Ω 上的一个集合代数,再设$\succeq$是 $\mathfrak{I}^*$ 上的一个二元关系,$\mathfrak{I}^*$ 是相对于 $\mathfrak{I}$ 的扩展的指标函数的代数。于是,定性代数$(\Omega, \mathfrak{I}, \succeq)$在定性的意义上是可满足的,当且仅当,对于 $\mathfrak{I}^*$ 中的每个 $\boldsymbol{A}^*$、$\boldsymbol{B}^*$ 和 $\boldsymbol{C}^*$,满足

下列公理：

公理 1. 关系$\succeq$是$\mathfrak{I}^*$的一个弱排序；

公理 2. $\boldsymbol{\Omega}^i \succ \varnothing^i$；

232 公理 3. $\boldsymbol{A}^* \succeq \varnothing^i$；

公理 4. $\boldsymbol{A}^* \succeq \boldsymbol{B}^*$，当且仅当，$\boldsymbol{A}^* + \boldsymbol{C}^* \succeq \boldsymbol{B}^* + \boldsymbol{C}^*$；

公理 5. 如果 $\boldsymbol{A}^* \succeq \boldsymbol{B}^*$，那么，对于在$\mathfrak{I}$中的每个$\boldsymbol{C}^*$和$\boldsymbol{D}^*$，存在着一个正整数 n，使得

$$n\boldsymbol{A}^* + \boldsymbol{C}^* \succeq n\boldsymbol{B}^* + \boldsymbol{D}^*$$

从有关定性概率的文献来看，这些公理似乎应该很熟悉。注意，公理 4 是加法公理，非常类似于德·菲内蒂的事件的加法公理：如果 $A \cap C = B \cap C = \varnothing$，那么 $A \succeq B$，当且仅当，$A \cup C \succeq B \cup C$。当我们从事件转向扩展的指标函数时，函数相加取代了并集。在公理 4 的精确阐述中已经看到了这种措施在形式上的重要性。扩展的指标函数的加法是无条件的——没有限于对应于 $A \cap C = B \cap C = \varnothing$。缺少这种限制有着深远的形式结果，它允许我们在没有任何真正的修改的情况下运用一般的扩展测量理论。公理 5 事实上是阿基米德公理的精确形式，克兰茨等人(1971，第 73 页)在给出扩展测量的充分必要条件时运用了这个公理。

定理 4. 设 Ω 是一个非空集，设$\mathfrak{I}$是 Ω 上的一个集合代数，再设$\succeq$是$\mathfrak{I}$上的一个二元关系。于是，在$\mathfrak{I}$上存在着一个严格一致的概率测度的充分必要条件是，存在着从$\mathfrak{I}$到$\mathfrak{I}^*$的扩展，使得扩展的指标函数$(\Omega, \mathfrak{I}^*, \succeq)$的定性代数在定性的意义上是可满足的。此外，如果$(\Omega, \mathfrak{I}^*, \succeq)$在定性的意义上是可满足的，那么在$\mathfrak{I}^*$上惟一地存在着一个严格一致的期望函数，而且，这个期望函数在$\mathfrak{I}^*$上惟一地产生了一个严格一致的概率

测度。

证明：正如已经表明的那样，在这个证明中所用的主要结果来自扩展的测量理论：存在着数值表征的充分必要条件。特别是，设 A 是一个非空集，$\succeq$ 是 A 上的一个二元关系，$\circ$ 是闭合的 A 上的一个二元运算，于是，在 A 上存在着一个数值函数，惟一地等于一个正的相似变换(即，乘以一个正的实数)，使得对于在 A 中的 a 和 b，

(i) $\phi(a) \geqslant \phi(b)$，当且仅当，$a \succeq b$，

(ii) $\phi(a \circ b) = \phi(a) + \phi(b)$

当且仅当，对于 A 中所有的 a、b、c 和 d，都满足下列四个公理：

E1. 关系 $\succeq$ 是 A 的一个弱排序；

E2. $a \circ (b \circ c) \approx (a \circ b) \circ c$，这里，$\approx$ 是根据 $\succeq$ 定义的等价关系；

E3. $a \succeq b$，当且仅当，$a \circ c \succeq b \circ c$，当且仅当，$c \circ a \succeq c \circ b$；

E4. 如果 $a \succ b$，那么，对于在 A 中的任何一个 c 和 d 来说，存在着一个正整数 n，使得 $na \circ c \succeq nb \circ d$，这里，在归纳意义上定义 na。

很容易检查，像上面定义的那样，在定性意义上令人满意的扩展的指标函数的代数，满足扩展测量结构的这四个公理。首先，我们注意到，函数相加在 $\mathfrak{J}^*$ 上是闭合的。第二，公理 1 相当于 E1。扩展的公理 E2 是直接从数值函数相加的结合性得出的，也就是说，对于 $\mathfrak{J}^*$ 中的任何一个 $\boldsymbol{A}^*$、$\boldsymbol{B}^*$ 和 $\boldsymbol{C}^*$， 233

$$\boldsymbol{A}^* + (\boldsymbol{B}^* + \boldsymbol{C}^*) = (\boldsymbol{A}^* + \boldsymbol{B}^*) + \boldsymbol{C}^*$$

因此，我们恰好没有等价，而是有恒等。公理 E3 是从公理 4 和数值函数相加是可交换的这个事实得到的。最后，E4 是从基本上完全相同的公理 5 得到的。

因此,对于定性地令人满意的任何一个代数(Ω, $\mathfrak{I}^*$, $\succeq$)来说,我们能够推出,存在着在 $\mathfrak{I}^*$ 上的一个数值函数 ϕ,使得对于在 $\mathfrak{I}^*$ 中的 $\boldsymbol{A}^*$ 和 $\boldsymbol{B}^*$,

(i) $\phi(\boldsymbol{A}^*) \geqslant \phi(\boldsymbol{B}^*)$,当且仅当,$\boldsymbol{A}^* \succeq \boldsymbol{B}^*$

(ii) $\phi(\boldsymbol{A}^* + \boldsymbol{B}^*) = \phi(\boldsymbol{A}^*) + \phi(\boldsymbol{B}^*)$

(iii) 最小的$(\boldsymbol{A}^*) \leqslant \varphi(\boldsymbol{A}^*) \leqslant$ 最大的$(\boldsymbol{A}^*)$

因为对于一个 $\boldsymbol{A}^*$,最小和最大都存在。

第二,因为对于在 $\mathfrak{I}^*$ 中的每个 $\boldsymbol{A}^*$,

$$\boldsymbol{A}^* + \varnothing^i = \boldsymbol{A}^*$$

所以,我们从(ii)立即有

$$\phi(\varnothing^i) = 0$$

因为 $\boldsymbol{\Omega}^i$ 同样是 1,因此,满足(iii)

$$\phi(\boldsymbol{\Omega}^i) = 1$$

这样,对于扩展的指标函数来说,ϕ 是一个标准的(惟一的)期望函数 E:

(i) $E(\varnothing^i) = 0$

(ii) $E(\boldsymbol{\Omega}^i) = 1$

(iii) $E(\boldsymbol{A}^* + \boldsymbol{B}^*) = E(\boldsymbol{A}^*) + E(\boldsymbol{B}^*)$

但是,当把 $\mathfrak{I}^*$ 的这样一个期望函数限于 $\mathfrak{I}^*$ 中的指标函数时,这定义了一个在 $\mathfrak{I}$ 上惟一的概率测度 P,即对于 $\mathfrak{I}$ 中的 A,我们定义

$$P(A) = E(\boldsymbol{A}^i)$$

因此,这些公理是充分的,但同样明显的是,只有公理 2 和公理 3 超越了扩展结构,也是概率表征所需要那些公理。从扩展的指标函数的特征来看,对于每个概率测度来说,都存在着从 $\mathfrak{I}$

惟一地扩展到 $\mathfrak{F}^*$ 的定性排序,这也是明确的。

证毕。

刚才给出的这个证明,甚至比陈述定理本身,更表明了在一个概率空间上定义什么样的随机变量的子集,才能以一种自然的方式足以确定概率测度。我们的步骤是,以定性方式使扩展的指标函数的期望公理化。根本没有必要考虑所有的随机变量,另一方面,指标函数的更加受限制的集合,在处理事件代数时,产生了同样的公理化困难。

定性的条件概率。定性的条件概率理论更加令人困惑的问题之一是,$A|B$ 不是一个对象——特别是,它不是以某种方式由事件 A 和 B 组成的一个新事件。因此,定性的理论依赖于一个四元关系 $A \mid B \succeq C \mid D$,把这个关系读作:已知事件 B 出现事件 A 的可能,至少与已知事件 D 出现事件 C 的可能一样大。使这种四元关系公理化的企图有许多(Koopman, 1940a, b; Aczél, 1961, 1966, p. 319; Luce, 1968; Domotor, 1969; Krantz et al., 1971; and Suppes, 1973a)。讨论给出充分必要条件问题的这些公理化中只有一种公理化是德默特尔(Domotor)的工作,他的工作接近于有穷情况下的公理化问题,在风格上类似于斯科特(1964)的公理化问题。 234

通过运用指标函数,或者,更一般地说,通过运用扩展的指标函数,排除了 $A|B$ 不是一个代数的困难,因为 $\boldsymbol{A}^i|B$ 恰好是集合 A 限于集合 B 的指标函数,即 $\boldsymbol{A}^i|B$ 是一个偏函数,它的定义域是 B。以同样的方式,如果 $\boldsymbol{X}$ 是一个扩展的指标函数,$\boldsymbol{X}|A$ 就是限于集合 A 的函数。在阐述我们感兴趣的函数代数时,需要注意这些偏函数的使用,因为当 $A \neq B$,但是 $A \cap B \neq \varnothing$ 时,将没有很好地定义函数的加法 $\boldsymbol{X} \mid A + \boldsymbol{Y} \mid B$。因此,为了十分明确,像以前一样,我们从一个非空集 Ω,即概率

空间和一个事件代数 $\mathfrak{I}$ 开始,即从 Ω 的子集开始,据此理解,$\mathfrak{I}$ 在并集和补集条件下是闭合的。接下来,我们把这个代数延伸到扩展的指标函数的代数 $\mathfrak{I}^*$,即包括在 $\mathfrak{I}$ 中的所有事件的指标函数在内的最小半群(在函数条件下)。现在,把后面这个代数扩展到也包括 Ω 上的所有偏函数在内,Ω 上偏函数是限于 $\mathfrak{I}$ 中的一个事件的扩展的指标函数。我们把这种代数称为偏扩展的指标函数 $\mathfrak{R}\mathfrak{I}^*$ 或 $\mathfrak{R}\mathfrak{I}^*(\Omega)$(如果需要完全明确的话)。从这个定义来看,显然,如果 $\boldsymbol{X}|A$ 和 $\boldsymbol{Y}|B$ 在 $\mathfrak{R}\mathfrak{I}^*$ 中,那么①

如果 $A = B$, $\boldsymbol{X} \mid A + \boldsymbol{Y} \mid B$ 在 $\mathfrak{R}\mathfrak{I}^*$ 中;

如果 $A \cap B = \varnothing$, $\boldsymbol{X} \mid A \cup \boldsymbol{Y} \mid B$ 在 $\mathfrak{R}\mathfrak{I}^*$ 中。

在决策理论或期望效用理论的更一般背景中,为了直接比较一个人对不同限制领域的两个决策函数的期望或偏爱,对他的直观能力的讨论相当多。我们没有评论这些文献,但确实想声明,我们在进行这样的比较时,没有发现明显的困难。个别情况可能会出现一些问题,但并不一定,因为定义域不同。事实上,我们认为,在不同条件下比较期望值是日常经验中很熟悉的一个问题。在当前的背景下,定性地比较所限制的期望值,可能被认为只是研究信念,而不是研究效用。基本的排序关系是 $\mathfrak{R}\mathfrak{I}^*$ 的一个弱排序,以标准方式定义的严格排序 $\succ$ 和等价关系 $\approx$。

235 沿着苏佩斯和扎诺蒂(1982)的思路,当要求涵盖随机变量的期望时,我们提供了这样的公理:这些公理强到足以证明所构造的概率测度是惟一的。值得对这种惟一性问题作出更多的

---

① 在这一小节,为了避免在 $\boldsymbol{X}|A$ 等条件下的任何混淆,把扩展的指标函数的符号 $\boldsymbol{A}^*$ 和 $\boldsymbol{B}^*$ 改变为 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$。

讨论。所提到的早期的文章全部集中于概率分布的存在性，但是，从一个令人满意的理论的观点来看，人们希望一种惟一分布，显然有许多不同的理由。例如，如果我们超越排序特性，只有相当于分布的凸多面体的惟一性，比如，斯科特的有穷概率空间的定理就是这种情况，那么我们就不能以一种自然的方式研究复合假设，因为概率加法是没有意义的。

定义 3. 设 Ω 是一个非空集，设 $\mathfrak{R}\mathfrak{J}^*(\Omega)$ 是偏扩展的指标函数的一个代数，再设 $\succeq$ 是 $\mathfrak{R}\mathfrak{J}^*$ 上的一个二元关系。于是，结构 $(\Omega, \mathfrak{R}\mathfrak{J}^*, \succeq)$ 是一个部分定性的期望结构，当且仅当，对于 $\mathfrak{J}^*$ 中的每个 $\boldsymbol{X}$、$\boldsymbol{Y}$、$\boldsymbol{X}_1$、$\boldsymbol{X}_2$、$\boldsymbol{Y}_1$ 和 $\boldsymbol{Y}_2$，以及 $\mathfrak{J}$ 中具有 $A, B \succ \varnothing$ 的每个 A、B 和 C，都满足下列公理：

公理 1. 关系 $\succeq$ 是 $\mathfrak{R}\mathfrak{J}^*$ 的一个弱排序；

公理 2. $\boldsymbol{\Omega}^i \succeq \varnothing^i$；

公理 3. $\boldsymbol{\Omega}^i \mid A \succeq \boldsymbol{C}^i \mid B \succeq \varnothing^i \mid A$；

公理 4a. 如果 $\boldsymbol{X}_1 \mid A \succeq \boldsymbol{Y}_1 \mid B$ 和 $\boldsymbol{X}_2 \mid A \succeq \boldsymbol{Y}_2 \mid B$，那么

$$\boldsymbol{X}_1 \mid A + \boldsymbol{X}_2 \mid A \succeq \boldsymbol{Y}_1 \mid B + \boldsymbol{Y}_2 \mid B$$

公理 4b. 如果 $\boldsymbol{X}_1 \mid A \preceq \boldsymbol{Y}_1 \mid B$ 和 $\boldsymbol{X}_1 \mid A + \boldsymbol{X}_2 \mid A \succeq \boldsymbol{Y}_1 \mid B + \boldsymbol{Y}_2 \mid B$，那么

$$\boldsymbol{X}_2 \mid A \succeq \boldsymbol{Y}_2 \mid B$$

公理 5. 如果 $A \subseteq B$，那么

$$\boldsymbol{X} \mid A \succeq \boldsymbol{Y} \mid A \text{，当且仅当，} \boldsymbol{X} \cdot \boldsymbol{A}^i \mid B \succeq \boldsymbol{Y} \cdot \boldsymbol{A}^i \mid B$$

公理 6.（阿基米德公理）。如果 $\boldsymbol{X} \mid A \succ \boldsymbol{Y} \mid B$，那么，对于 $\mathfrak{J}^*$ 中的每个 Z 来说，存在着一个正整数 n，使得

$$n\boldsymbol{X} \mid A \succeq n\boldsymbol{Y} \mid B + \boldsymbol{Z} \mid B$$

这些公理在特征上是简单的，而且，它们与定义 2 的公理的关系是明显的。前三个公理非常简单。公理 4，即加法公理，必须是被相对地应用于所限制的集合。注意，我们在这个不等式的两边有不同的限制。

我无法表明，是否有可能用下列较弱的和更自然的公理来取代公理 4 的两个部分。如果 $\boldsymbol{X}_2 \mid A \approx \boldsymbol{Y}_2 \mid B$，那么 $\boldsymbol{X}_1 \mid A \succeq \boldsymbol{Y}_1 \mid B$，当且仅当，$\boldsymbol{X}_1 \mid A + \boldsymbol{X}_2 \mid A \succeq \boldsymbol{Y}_1 \mid B + \boldsymbol{Y}_2 \mid B$。实际上新的公理是公理 5。从事件和数值概率的角度来看，这个公理对应于下列公理：如果 $A \subseteq B$，那么 $P(C \mid A) \geqslant P(D \mid A)$，当且仅当，$P(C \cap A \mid B) \geqslant P(D \cap A \mid B)$。注意，在这个公理本身，函数的乘法运算取代了事件的交集。(在函数的乘法运算的条件下，容易证明 $\mathfrak{J}^*$ 是闭合的。)公理 6 是熟悉的和必要的阿基米德公理。

现在，我们陈述和证明主要定理。在这个定理中，我们谈到在 $\mathfrak{R}\mathfrak{J}^*(\Omega)$上的一个严格一致的期望函数。从标准概率理论和条件期望效用理论来看，这种期望值的特性显然应该如下，对于 $A,\ B \succ \varnothing$， 236

(i) $E(\boldsymbol{X} \mid A) \geqslant E(\boldsymbol{Y} \mid B)$，当且仅当，$\boldsymbol{X} \mid A \succeq \boldsymbol{Y} \mid B$；

(ii) $E(\boldsymbol{X} \mid A + \boldsymbol{Y} \mid A) = E(\boldsymbol{X} \mid A) + E(\boldsymbol{Y} \mid A)$；

(iii) $E(\boldsymbol{X} \cdot \boldsymbol{A}^i \mid B) = E(\boldsymbol{X} \mid A)E(\boldsymbol{A}^i \mid B)$，如果 $A \subseteq B$；

(iv) $E(\varnothing^i \mid A) = 0$ 和 $E(\boldsymbol{\Omega}^i \mid A) = 1$。

主要运用(iii)，就容易证明，在前面提到的公理化文献中出现的下列特性：

$$E(\boldsymbol{X} \mid A \cup \boldsymbol{Y} \mid B) = E(\boldsymbol{X} \mid A)E(\boldsymbol{A}^i \mid A \cup B) + E(\boldsymbol{Y} \mid B)E(\boldsymbol{B}^i \mid A \cup B)$$

因为 $A \cap B = \varnothing$

定理 5. 设 Ω 是一个非空集，设 $\mathfrak{J}$ 是 Ω 上的一个集合代数，再设 $\geq$ 是 $\mathfrak{J}\times\mathfrak{J}$ 上的一个二元关系。于是，存在着在 $\mathfrak{J}\times\mathfrak{J}$ 上的一个严格一致的条件概率测度的充分必要条件是，存在着 $\geq$ 的从 $\mathfrak{J}\times\mathfrak{J}$ 到 $\mathfrak{R}\mathfrak{J}^*(\Omega)$ 的一个扩展 $\geq^*$，使得结构 $(\Omega, \mathfrak{R}\mathfrak{J}^*(\Omega), \geq^*)$ 是一个部分定性的期望结构。此外，如果 $(\Omega, \mathfrak{R}\mathfrak{J}^*(\Omega), \geq^*)$ 是一个部分定性的期望结构，那么存在着在 $\mathfrak{R}\mathfrak{J}^*(\Omega)$ 上的一个惟一严格一致的条件期望函数，而且，这个期望值产生了在 $\mathfrak{J}\times\mathfrak{J}$ 上的一个惟一严格一致的条件概率测度。

在苏佩斯和扎诺蒂的文章中(1982)给出了证明，我这里采纳了这个证明。

证明： 对于具有 $A \succ \varnothing$ 的每个 $\boldsymbol{X}|A$ 来说，我们定义集合

$$S(\boldsymbol{X} \mid A) = \left\{\frac{m}{n}: m\boldsymbol{\Omega}^i \mid A \geq n\boldsymbol{X} \mid A\right\}$$

(我们注意到，从这些公理很容易证明 $\boldsymbol{\Omega}^i \approx \boldsymbol{\Omega}^i \mid A$，因此，为了达到一般的目标，我们能够写出：$m\boldsymbol{\Omega}^i \geq n\boldsymbol{X} \mid A$。）已知这个定义，基于在公理 1—4 和 6 的定理 4 的证明中简化为已知的扩展测度的充分必要条件(克兰茨等人，1971，第 3 章)，我们首先知道，存在 $S(\boldsymbol{X}|A)$ 的下限。根据克兰茨等人的证明，我们用此来定义已知 A 时的 $\boldsymbol{X}$ 的期望值：

$$E(\boldsymbol{X} \mid A) = g.\,l.\,b. \left\{\frac{m}{n}: m\boldsymbol{\Omega}^i \geq n\boldsymbol{X} \mid A\right\} \tag{3}$$

于是，从前面的那些结果得出，函数 E(对于固定的 A 来说)是惟一的，而且：

$$E(\boldsymbol{X} \mid A) \geqslant E(\boldsymbol{Y} \mid A)\text{，当且仅当，}\boldsymbol{X} \mid A \geq \boldsymbol{Y} \mid A \tag{4}$$

$$E(\boldsymbol{X} \mid A + \boldsymbol{Y} \mid A) = E(\boldsymbol{X} \mid A) + E(\boldsymbol{Y} \mid A) \tag{5}$$

$$E(\varnothing^i \mid A) = 0 \text{ 和 } E(\boldsymbol{\Omega}^i \mid A) = 1 \tag{6}$$

现在关键的一步是,把这个结果扩展到已知事件 A 和 B 之间的关系。

我们首先证明,通过期望函数保持排序。对于这个证明的前半部分来说,假设

$$\mathbf{X} \mid A \succeq \mathbf{Y} \mid B \tag{7}$$

相反,再假定

$$E(\mathbf{Y} \mid B) > E(\mathbf{X} \mid A) \tag{8}$$

237 于是,一定存在着自然数 m 和 n,使得

$$E(\mathbf{Y} \mid B) > \frac{m}{n} > E(\mathbf{X} \mid A) \tag{9}$$

因此,从函数 E 的定义来看,我们有

$$m\boldsymbol{\Omega}^i \prec n\mathbf{Y} \mid B \tag{10}$$

和

$$m\boldsymbol{\Omega}^i \succeq n\mathbf{X} \mid A \tag{11}$$

由此

$$n\mathbf{Y} \mid B \succ n\mathbf{X} \mid A \tag{12}$$

但从(7)和公理 4a 来看,我通过一个简单的归纳得出

$$n\mathbf{X} \mid A \succeq n\mathbf{Y} \mid B \tag{13}$$

这与(12)相矛盾,因此,假定(8)是错误的。

现在假设

$$E(\mathbf{X} \mid A) \geqslant E(\mathbf{Y} \mid B) \tag{14}$$

再假定

$$\mathbf{Y} \mid B \succ \mathbf{X} \mid A \tag{15}$$

现在,如果 $E(\mathbf{X} \mid A) > E(\mathbf{Y} \mid B)$,通过刚才给出的这种论证,我

们能立即表明

$$\mathbf{X} \mid A \succ \mathbf{Y} \mid B \tag{16}$$

这与(15)相矛盾,另一方面,如果

$$E(\mathbf{X} \mid A) = E(\mathbf{Y} \mid B) \tag{17}$$

那么我们能作出下列论证。根据(15)和公理 6,存在着一个 n,使得

$$n\mathbf{Y} \mid B \succeq (n+1)\mathbf{X} \mid A \tag{18}$$

由此,根据前面的论证

$$E(n\mathbf{Y} \mid B) \geqslant E((n+1)\mathbf{X} \mid A) \tag{19}$$

而且根据(5)

$$nE(\mathbf{Y} \mid B) \geqslant (n+1)E(\mathbf{X} \mid A), \tag{20}$$

因此,根据(17)和(20)

$$E(\mathbf{Y} \mid B) \leqslant 0 \tag{21}$$

但从(4)—(6)来看,容易得出

$$E(\mathbf{Y} \mid B) \geqslant 0 \tag{22}$$

由此

$$E(\mathbf{Y} \mid B) = 0 \tag{23}$$

然后,再一次运用(4)—(6),我们获得

$$\mathbf{Y} \mid B \approx \varnothing^{i} \mid B \tag{24}$$

而且借助于公理 3,

$$\mathbf{X} \mid A \succeq \varnothing^{i} \mid B \tag{25}$$

238 由此，根据传递性，从(24)和(25)得出

$$\mathbf{X} \mid A \succeq \mathbf{Y} \mid B \tag{26}$$

与(15)相矛盾。因此我们现在已经表明

$$E(\mathbf{X} \mid A) \geqslant E(\mathbf{Y} \mid B)\text{，当且仅当，}\mathbf{X} \mid A \succeq \mathbf{Y} \mid B \tag{27}$$

最后，我们需要证明，对于 $A \succ \varnothing$ 和 $A \subseteq B$，

$$E(\mathbf{X} \cdot \mathbf{A}^i \mid B) = E(\mathbf{X} \mid A) E(\mathbf{A}^i \mid B) \tag{28}$$

我们首先注意到，通过在公理 5 中用 $m\boldsymbol{\Omega}^i$ 表示 $\mathbf{X}$，用 $n\mathbf{X}$ 表示 $\mathbf{Y}$，我们得到

$$m\boldsymbol{\Omega}^i \geqslant n\mathbf{X} \mid A\text{，当且仅当，}m\mathbf{A}^i \mid B \geqslant n\mathbf{X} \cdot \mathbf{A}^i \mid B \tag{29}$$

从(29)直接得出

$$\left\{\frac{m}{n}\text{: } m\boldsymbol{\Omega}^i \succeq n\mathbf{X} \mid A\right\} = \left\{\frac{m}{n}\text{: } m\mathbf{A}^i \mid B \succeq n\mathbf{X} \cdot \mathbf{A}^i \mid B\right\} \tag{30}$$

由此，它们的下限是一样的，而且，我们有

$$E(\mathbf{X} \mid A) = E'_{\mathbf{A}^i \mid B}(\mathbf{X} \cdot \mathbf{A}^i \mid B) \tag{31}$$

这里，E' 是将 $\mathbf{A}^i \mid B$ 作为一个单元的测度函数，也就是说，

$$E'_{\mathbf{A}^i \mid B}(\mathbf{A}^i \mid B) = 1$$

像在扩展测度理论中常见的那样，存在一个正实数 c，使得对于每个 $\mathbf{X}$，

$$cE'_{\mathbf{A}^i \mid B}(\mathbf{X} \cdot \mathbf{A}^i \mid B) = E(\mathbf{X} \cdot \mathbf{A}^i \mid B) \tag{32}$$

现在通过(31)和取 $\mathbf{X} = \boldsymbol{\Omega}^i$，

$$cE(\boldsymbol{\Omega}^i \mid A) = E(\boldsymbol{\Omega}^i \cdot \mathbf{A}^i \mid B)$$

但是，$E(\boldsymbol{\Omega}^i \mid A) = 1$，因此，

$$c = E(\boldsymbol{\Omega}^i \cdot \mathbf{A}^i \mid B) = E(\mathbf{A}^i \mid B) \tag{33}$$

(31)、(32)和(33)结合起来，如期所料，我们获得(28)。

从(6)和前面关于无条件概率的定理 4 的结果，得到期望函数的惟一性。

对于 $A > \varnothing$，我们然后在 $\mathfrak{J}$ 中为每个 B 下定义，

$$P(B \mid A) = E(\boldsymbol{B}^i \mid A)$$

而且，表明函数 P 是 $\mathfrak{J}$ 上的一个条件概率测度(它确立了这些公理的充分性)，是小事一桩。很容易检查这些公理中每个公理的必要性。

定性公理的历史背景。正如已经表明的那样，德·菲内蒂(1937/1964)的定性公理不是第一套公理。[1] 凯恩斯(1921)和杰弗里斯与林奇(Wrinch)(1919)很早以前就在概率的逻辑理论框架内提供了这些公理的各种不完备集。无疑，在概率的主观理论框架内，历史上，德·菲内蒂的最重要的前身是拉姆齐(Ramsey)的"真理与概率"(1931)，这篇文章写于 1926 年，是他的遗作。拉姆齐的基本进路与德·菲内蒂的进路不同，他以行动和期望效用为出发点，或者，用现在熟悉的术语来说，是从信念和愿望出发，不是单独从信念出发。他在下列两段中很好地描述了这个一般纲领。

239

> 因此，为了建构一个既一般又更精确的信念量化理论，我建议把一般心理学理论当作基础，人们现在普遍地丢弃了这种理论，但我却认为，在我们最关心的那些情况下，它

① 在日常事务中定性地运用概率的非常详细的特别早的讨论，可以在休谟的《人性论》(1739)一书的不同章节中，特别是第 1 卷第 3 部分的第 13 节中找到。

相当地接近于真理。我指的这种理论是：我们以认为最有可能实现我们渴望的目标的方式来行动，因此，一个人的行动完全是由他的欲望和意见来决定的。这个理论不可能适当地符合所有的事实，但在我看来，特别是在我们的自我意识或职业生活的情况下，是对真理的有用的接近，而且，预设它就在我们的许多思想中。这是一个简单的理论，也是许多心理学家为了使它与事实更协调，明显地喜欢通过引入无意识的欲望和无意识的意见来坚持的理论。这样的幻想在多大程度上能达到所要求的结果，我不作判断：我只主张，追求近似真理，或者与这种人为的心理体系相关的真理，我认为，这就像牛顿力学的情况一样，即使知道它是错的，但仍然能够有效地运用。

我们必须看到，这个理论不能被列为是功利主义的心理学，在这个理论中，快乐占有统治地位。我建议采纳的这个理论是，我们寻找我们想要的东西，它们可以是我们自己或他人的快乐，或者，无论别的什么东西，而且，我们认为，我们的行动是最有可能实现这些美德的行动。但这不是一个精确的陈述，因为只有在我们引入信念量化概念之后，才能对这个理论作出精确的陈述。

(Ramsey, 1931: p. 173)

正如拉姆齐自己所言，他的思想只是在所引用的文章中得到了概述——关于这个主题他只写过这一篇文章。不过，这种概述是重要的，因为他勾画了一种简单明确的方式来发现，一个“在伦理意义上中立的命题 P 的可信度是 1/2”——他的第一公理。尽管他的公理只是随便阐述的，但是，他接下来引入了优先“世界”的一种排序关系，正如他指出的那样，通过衡量价值的进一

步的公理加以详述。于是，他建议，运用对价值的衡量，通过打赌比例的惯用方法，衡量信念的大小或概率。[1]

后来，戴维森(Davidson)和苏佩斯(1957)直接用拉姆齐的思想作为主观概率和效用的详细的有穷公理化的基础。再稍后，在所关注的实验中，用这些思想来衡量个人的概率和效用(Davidson, Suppes and Siegel, 1957)。

但是，不要忘记的基本观点是，与拉姆齐相反，德·菲内蒂在许多出版物中从数学上和哲学上很详尽地阐述了他的思想，拉姆齐在二十六岁去世后，突然中终止了任何进一步的工作。后来，德·菲内蒂产生了很大影响几乎不会令人感到惊奇。

德·菲内蒂的表征定理。像德·菲内蒂那样的主观主义者 240 的核心问题是，如何处理这样的标准(即使是人为的)问题：在公平地掷一枚可能有偏向的硬币时，如何估计出现正面的客观上未知的概率 p。相对频率进路恰好是从一个无穷试验序列估计 p。(人为性来自使用一个无穷序列，在有限时间内，产生或看到这个序列是不可能的。)德·菲内蒂的回应是，假设同样人为地使用一个无穷序列，对于主观主义者来说，这不是一个问题。[2] 已经几次提到，他所用的概念是可交换性概念，我现在用无穷序列来定义这个概念。在一个概率空间上的随机变量 $\boldsymbol{X}_1$，$\boldsymbol{X}_2$，…，$\boldsymbol{X}_n$，…的一个 0，1 序列，是**可交换的**，如果对于每个 n 和前 n 个整数的每一种排列 σ，$\boldsymbol{X}_1$，$\boldsymbol{X}_2$，…，$\boldsymbol{X}_n$，…都有与 $\boldsymbol{X}_{\sigma 1}$，$\boldsymbol{X}_{\sigma 2}$，…，$\boldsymbol{X}_{\sigma n}$ 一样的概率分布。我在这个定理中需要陈述

---

① 德·菲内蒂在 1937 年的文章的英译本中增加了一个脚注，正如他在脚注中所说的那样(第 102 页)，他在 1937 年之前并不了解拉姆齐的工作，他的这篇文章是根据他两年前在巴黎的演讲写成的。

② 有穷序列的可交换性的明显结果可以在迪亚科尼斯(Diaconis, 1977)以及迪亚科尼斯与弗里德曼(Freedman, 1980)的文章中找到。

的另一个概念是混合概率的概念。这里我们稍微简化一些，把混合概率写成在区间[0, 1]上的概率密度 f。因此，f 具有特性：

(i) 对于在[0, 1]中的 p, $f(p) \geqslant 0$

(ii) $\int_0^1 f(x)\mathrm{d}x = 1$

现在，很容易表明，独立序列 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n, \cdots$的任何一种混合 f 都是可交换的。德·菲内蒂证明了最难的部分，即逆命题。

定理 6. (de Finetti 1937/1964)只取值为 0 和 1 的随机变量 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n, \cdots$的一个序列，是可交换的，当且仅当，存在着在[0, 1]上的一个混合密度 f，使得进行 n 次试验有 $m(0 \leqslant m \leqslant n)$ 次成功的任何一个样本的概率，即这个样本的概率是

$$\int_0^1 p^m(1-p)^{n-m}f(p)\mathrm{d}p$$

另外一个引人注意的阐述是，在 n 次试验的一个样本中，看到正反面的有穷序列的概率，只依赖于两个数，样本长度 n 和出现正面的个数 m。这是对应于第一节最后给出的表述的一个不变性的陈述。还要注意的最重要事实是，已知样本，结果 p 是独立的。德·菲内蒂(1937/1964)对这个定理的认真而详细的哲学讨论，是非常值得推荐的。德·菲内蒂还详细地提供了证明，这个证明也可推广到随机量，用德·菲内蒂的术语来说，是一般的随机变量。我最好用德·菲内蒂自己总结的一段来结束这里的讨论：

> 我们已经达到的结果为我们提供了所寻找的答案，这个答案是非常简单的和很令人满意的：应该用“可交换事件”的定义替代“确定但未知概率的独立事件”的含糊的和

不能令人满意的定义。这种回答提供了直接应用于估计单个事件概率的一个条件,而且,不会遇到主观主义的考虑建议排除的任何一个困难。它构成了还原到下列要求的很自然和很明确的一个条件,这个条件具有纯定性的特征:这个要求是,对特定事件作出等概率的判断,或更确切地说, 241
n 个事件 E_{i_1} ,…, E_{i_n} 的所有结合,总有相同的概率,无论 E_i 的选择还是排序如何。与我们的概率判断相关的同样简单的"对称性"条件定义了可交换的随机量,并且在一般情况下定义了任何空间中的可交换的随机元素。这在所有情况下都导致了同样切实可行的结论:足够丰富的经验使我们总是认为,未来可能的频率或分布接近于已看到的那些频率或分布。

(de Finetti, 1937/1964: p. 142)

客观先验的辩护。客观先验概率的概念已经出现在第5节以几种不同形式讨论的概率的逻辑理论中。但还有另外一个看法,有时称之为一种客观的贝叶斯主义的观点,这种观点坚持认为,只应该使用基于预先可靠的经验结果和科学结果的客观先验概率。关于这一点的最直接的进路,是一位物理学家杰恩斯(E. T. Jaynes)提供的进路,他的主要文章被收集在罗格·罗森克兰茨(Roger Rosenkrantz)主编的文集中(1983)。杰恩斯在许多文章中强烈地论证了用最大熵原理作为这种客观先验。特别是在杰恩斯1978年的文章中。他表明,这个原理如何肯定在统计力学中具有意义,而且,如何致使直接推导出麦克斯韦、玻尔兹曼(Boltcmann)和吉布斯的经典结果。

如果我们限于在一个有穷集上的一个离散密度,最大熵原理就是选择客观先验概率,以使这种分布的熵 H 极大化,即已

知受到的所有客观限制，使

$$H = -\sum_{i=1}^{n} p_i \log p_i$$

最大化。标准的论证是，这种分布事先是最不确定的。在4.5节中概述了，对于伯努利或马尔可夫过程来说，熵是完全不变的证明，有助于支持这种主张，尽管在那一节总结的各态历经结果，没有普遍地在客观先验的语境中加以讨论。①

242 第二种论证来自熟悉的统计力学。最大熵分布是所选择的分布，因为除了那些已经说明的限制之外，再没有其他限制，在这个零假设条件下，大多数可能的分布将接近于最大熵分布。下面是来自杰恩斯(1979)的一个熟悉的例子，这个例子举例说明了这一点，也表明了这个原理是如何起作用的。

假设把一枚骰子投掷一千次。如果"真的"掷骰子，预计朝

---

① 这种帮助来源于下列形式。使熵极大化，就是使不确定性极大化，至少对于马尔可夫过程来说是如此——伯努利过程是马尔可夫过程的一种特殊情况，因为对于这样的过程来说，具有相同熵的那些过程，有同构的不确定结构，而且，像在4.5节中说明的那样，反之亦然(原文"conversly" 有误，应该是"conversely"——译者)。至于熵为什么很好地衡量了不确定性，对这个问题的更熟悉的标准论证，参见Rosenkrantz(1977, p. 13ff)。我引用他的第13页的总结。"熵是对不确定性的一种令人满意的衡量，这一点通过熵的下列特性得到了进一步的证实：

(i) $H(p_1, \cdots, p_m) = H(p_1, \cdots, p_m, 0)$，熵完全是由赋予非零概率的替代选择来决定的。

(ii) 当所有的 p_i 都是相等的时，$H(p_1, \cdots, p_m)$ 是随着 m 的增加而增加的，m 是等概率选择的数目。

(iii) 当某个 $p_i = 1$ 时，$H(p_1, \cdots, p_m) = 0$，是极小值。

(iv) 当每个 $p_i = 1/m$ 时，$H(p_1, \cdots, p_m) = \log m$，是极大值。

(v) p_i 的任何一个平均值(即分布的任何一种扁平率)都使 H 增加。

(vi) H 是非负的。

(vii) 在标记 $1, \cdots, m$ 的每一种排列下，$H(p_1, \cdots, p_m)$都是不变的。

(viii) $H(p_1, \cdots, p_m)$在它自变量中是连续的。"

上的点数应该是 3.5。但是,我们被告知,对于稍微有偏向的骰子来说,正确的期望值是 4.5。因此,

$$\sum_{i=1}^{6} i f_i = 4.5$$

杰恩斯(1978)详细地推导出的结果是:$(f_1, \cdots, f_6) = (0.05, 0.08, 0.11, 0.17, 0.24, 0.35)$。偏向较多点的面是明显的。即使这个简单的例子,计算也是非平凡的,即在批评最大熵原理时,有时陈述的一种观点。对这种观点和其他观点的一个好的讨论,可以在罗森克兰茨的书(Rosenkrantz, 1977)中找到,这本书,特别是第 3 章,论述了这些观点和相关主题。

杰恩斯在许多物理事例的语境中最明确和最有效地辩护的这种客观贝叶斯主义观点,并没有胜过德·菲内蒂的个人主观主义。也许,德·菲内蒂强调的客观进路的主要困难,源于不适当地应用这些进路的企图。主观贝叶斯主义进路总是先验的,即使他无法说明或证明这一点。对于任何一个复杂而敏感的研究或应用领域来说,比如,天气预报,这是一种很得体的态度。但尽管如此,杰恩斯在统计力学中的某些应用是很有吸引力的,尽管它们可能是特殊的。此外,正如德·菲内蒂证实的那样,"客观性有它正确的、重要的方面"。①

一般论题。我现在转向在评价主观概率观时产生的许多问题。

对称性的使用。第一个自然的问题是问,主观理论如何利用这种对称性:它们是经典的拉普拉斯概率定义的一个自然的组成部分。如果我们根据贝叶斯定理进行思考,答案似乎是明

---

① 主要的引文取自德·菲内蒂(1975,第 201 页),但他在这一页和后面几页的冗长的哲学分析与归纳推理的概念问题很相关。

显的。我们在讨论赌博游戏时很自然地全部接受的对称性，直接被合并到作为先验概率的主观理论中。这样，例如，在一种牌戏中，已知亮出一张 A 的信息，如果我希望确定，下一轮发牌中获得 A 的概率，那么我用赌博游戏中的自然对称原理作为先验概率，赌博游戏是经典定义的一部分。当然，如果我希望提出精致的修正，我就能够这么做，特别是源于下列事实的修正：在通常洗牌的过程中，所引入的排列群是相对小的有穷阶的群，因此，尽管事实是，发牌人在发牌前已经相当彻底地洗过牌，但从一轮牌到另一轮牌仍然留有信息。关于洗牌的这些二阶改进，只标志着这种自然修正：这些修正是在主观理论中产生的，并且，在经典的概率定义框架内，原则上将是很难处理的。另一方面，我强调，在经典定义中所用的对称性原理是一位有经验的打牌者的先验概率的一个自然的组成部分。我们后面将返回的一种观点是，这些对称性在多大程度上是引人注目的。

243

相对频率的使用。同样应该明确的是，在计算后验概率时，是使用主观理论中的相对频率数据的正确地方。前面所举的应用贝叶斯定理的例子，提供了一个人为的但简单的事例，在这个事例中，这种用法是由相对频率数据构成的。为了在最初持不同意见的观察者之间达到趋同，需要什么是显而易见的。观察者必须在抽样方法上达成一致，当然，他们也必须对这种抽样的观察结果达成一致。在很弱的限定条件下，不管他们最初的意见有多么不同，他们在足够多的抽样观察的基础上，都能任意地接近于趋同。明显的必要条件是，个人的观察是近似于独立的，或至少是可交换的。例如，如果观察很受时间限制，那么许多观察只能算作是一次观察。既然已经考虑过这些问题，我们将不再详细地讨论它们。

对发生信念趋同的条件的反思，也阐明了不出现这种趋同

的许多情境。男女之间对宗教、经济和政治的根本不同的看法，就是缺乏趋同性领域的一个极好例子；无疑，这种分歧的一个主要来源，是没有在把什么算作是证据的问题上达成一致，或者，换句话说，不能对一个实验或收集数据的其他条件达成一致，它可以用一个达成一致可能性的函数来概括。

信息的强化特征问题。正如已经表明的那样，主观理论的一个重要方面是强调，同样有理性的人，对相同事件的概率，可能持有相当不同的观点。"**理性**"(rational)这个词的使用似乎超越了概率的主观理论所面对的问题。让我们考虑一两个例子说明如何能表达这种用法上的不同。

第一类例子是研究不同的人对信息的非语言处理。在预言明天的天气时，一个人总是比另一个人的预言更成功。至少在天气预报的有效力的数学方法——这种方法仅仅刚开始成为一种严肃的预报手段——问世之前，普遍看到的事实是，有经验的气象学家，即使受到共同的训练，具有相同的背景，面对一组共同的观察，在成功地预报世界的某个地区明天的天气时，他们的能力也有很大的不同。据我看来，根据标准的主观理论，例如，像德·菲内蒂所表达的主观理论，根本没有一种明确的方式来阐明，在一种场合，较好的预言家在处理他的信息时比其他人在某种意义上更理性；但在日常用法中，我们完全倾向于这样说。通常小说里的一段插曲和许多人的实际生活中的一个共同经历是，诋毁这样的人的智力或理性：他们面对许多相反的(即使也许是微妙的)证据，还天真地坚持相信他人的行为。 244

但是，连续的预测能够像任何其他的经验现象一样加以研究，而且，评价预报员的绩效的文献有很多，这是许多经验领域内的一个重要的实际话题。对评价主观和客观预报的定量方法的考察超出了本章的范围。《预测杂志》专门讨论这个问题。也

可参见,例如,马克里达基斯(Makridakis)等人(1984)和达维德(Dawid,1986)的工作。

与德·菲内蒂评论的许多大意相反,下列说法似乎是公平的:概率的主观理论提供了理性的必要条件,但没有提供理性的充分条件。

贝叶斯概率和概念形成问题。在如何处理信息的分析中,包括了反复考虑**错误**(mistaken)信念的一种重要观点。在日常用法中,当后来的信息表明一种信念是错误的时,这种信念通常被说成是搞错了或是非理性的。按照概率的主观理论和另外的更明智的习惯用法,这种错误信念的观点本身是一种错误。基于随后的而在信念被持有时并不存在的证据,没有表明这种信念是错误的。在根据新的信息从先验概率变化到后验概率时,反映了信念的适当转变。主观理论的重要观点是,总的概率测度本身是不变的,反而我们从先验的概率过渡到了后验的条件概率。我们所考虑的贝叶斯定理的应用足以证明这一点。下面来自德·菲内蒂的引文(1937/1964,第146页)很好地例示了这种观点。

> 在评论未来的预言时,无论受到什么影响,这都决不意味着,也决不意指,在概率 $P(E_{n+1})$ 被经验**否证**之后,我们**纠正**了最初估计的概率 $P(E_{n+1})$,并用符合这种经验因而有可能**更接近于真概率**的另一个概率 $P^*(E_{n+1})$ 取而代之;相反,这本身只在下列意义上体现出来:在开始的 n 次试验中,当经验向我们提供结果 A 时,我们的判断将不再通过概率 $P(E_{n+1})$ 来表达,而是通过概率 $P(E_{n+1}|A)$ 来表达,即我们最初的意见已经促成被考虑为以结果 A 为条件的事件 E_{n+1}。这个最初的意见没有被拒绝或纠正;它不是

> 已经修改了的函数 P(被另一个 P^* 取代),而是用 $E_{n+1}|A$ 取代了自变量 E_{n+1},而且,这恰好保留了我们对原始意见的信任(正如在选择函数 P 时所体现出来的那样)和在我们的下列判断中的一致性:当在已知的情况下发生一种变化时,我们的预言也随之变化。

尽管德·菲内蒂的论述是引人注目的,但是,似乎存在一大类情形,在这些情形中,应用他所确认的原理,是令人怀疑的。我想到了把一个真正的新概念强加于一个主题的所有这些情形。我不是意指,必定创造一个新的科学概念,而是指所有这样的情形:一个人在分析他面临的数据时,应用了从前没有用过的一个概念。① 245

未知概率的问题。概率的主观理论与概率概念的习惯用法相矛盾,它的另一个特征是这种观点:根本不存在未知概率。如果有人问我,明天斐济岛下雨的概率有多大,我的自然倾向是说,"我不知道",而不是试着给出一个概率估计。如果另一个人问我,我认为从现在开始的五百年内,斯坦福大学至少注册五万名学生的概率有多大,我自然倾向于只说,"我全然不知这个事件的概率或可能性有多大"。德·菲内蒂坚持的观点是,一个人总是有一种意见,因此,总有关于这些事情的一种概率估计,但似乎对我来说,这样一种观点根本没有内在的必要性。很容易看到它的一种来源。不管一个人对事件可能发生的情况所掌握的信息多么贫乏,他对任何一个事件总是有一种概率估计,这种

---

① I. J. 古德(Good,1983)在"演化的概率"的标题下标示的许多段落中,很好地强调了这种观点。甚至一个更重要的观点是,把新的信息条件化,正如德·菲内蒂强调的那样,尽管这可能是基本的,但不是取代明确发展贝叶斯统计的需求。一个极好和完全可理解的介绍来自林德利(Lindley,1971)。

要求起因于对二值逻辑的一个直接延伸。任何一个陈述要么是真的，要么是假的，相应地，针对每个被质问的人，任何一个陈述或事件，必须总有一个明确的概率。从一种形式的观点来看，似乎难处理的是，拥有由 0 和 1 之间的任何实数构成的一种逻辑，再加上毫不相关的值："我不知道"。

稍后我们将考察这种观点：一个人只有很少关于事件的背景信息，人们通过问他，就此事他下什么样的赌注，总能得出事件的主观概率。我没有预先考虑这种讨论，但仍然喜欢坚持认为，实际上，要求把一个概率赋予每个虚构的事件，似乎不是主观理论的一个固有的组成部分。以同样的精神，我们通过说我们根本不知道，回答了关于一个陈述的真理性的问题，我们也能以同样的方式，答复估计概率的请求。

决策和主观概率的测度。一个人的行动或决策，而不是他的言语，是他的信念的真标志，这种评论是老生常谈。当对这种普遍公认的原理作出反思时，就出现了大量讨论：一个人基于实际上做出的决策，如何测度主观概率。这是一个复杂的主题，我不试图对此作出很详细的论述。

246 根据主观概率的经典回应是，我们通过要求一个人打赌来发现他真正持有的主观概率。例如，如果他认为明天下雨的概率是 1/2，那么他将愿意在这个事件上下同额赌注。如果他认为，明天下雪的概率有 0.1 的主观概率，那么他将以 1∶9 的比率打赌不下雪。同样清楚的是，如何可以用这个同样的方法检验精确的陈述。例如，如果一个人说，明天下雨的概率至少是 1/2，那么，可以推测，任何一种赌注，只要提供了至少对他有利的可能性，他都将接受。

不幸的是，测度主观概率的这种方法，存在着一个核心的困难。如果赌注的金额发生变化，几乎所有的人都会改变他们将

同意与别人打赌的可能性。例如,一个人基于明天将要下雨的概率,接受同额赌注的打赌,他的理解是,如果事实上明天确实会下雨,他将会赢,如果赌注金额从一方的一美元增加到另一方的一百美元,他将不接受同额赌注的打赌。随便赌一美元的人,在相同情况下和在相同可能的条件下,将不愿意赌一百美元,肯定不会赌一千美元。在明天下雨的问题上接受同额赌注的人,只要赌注的等级是一百美元,也许愿意接受他能以二比一获胜的可能性,而当赌注金额更大时,他将只接受还要更有利的可能性。如果我们用这种下赌注的方法打赌,那么,我们说,他真正估计到的明天下雨的概率是多少呢?

尽管有这种批评,但用我们刚才描述的简单框架来测度主观概率,一直有许多有趣的经验研究。短暂地偏离来对其中的某些经验结果作出描述,可能是令人感兴趣的。再一次,我沿着卢斯和苏佩斯(1965)文章的思路加以讨论。

普雷斯顿(Preston)和巴拉塔(Baratta, 1948)很显然首先试图根据实验测度主观概率。受试者大体是由两人或多人组成的几个小组,他们常常在一个简单的拍卖游戏中下赌注竞标。在竞标结束后,成功的竞标人被允许掷一副骰子。掷骰子获胜的概率恰好对应于拍卖卡上规定的概率。例如,在一次已知的打赌中,受试者可能以 0.25 的获胜概率,竞标 250 点的奖金。如果就这次打赌而言,平均成功的出价是 50,那么庄家计算出的心理概率是 50/250＝0.20。运用这种计算方法,也许,他们最感兴趣的结论是,系统地高估了小于 0.20 的客观概率,而系统地低估了大于 0.20 的客观概率。

两个其他结论是值得注意的。第一,通过对老练的受试者的数据与对概率事实相对无知的那些受试者的数据作比较,他们断定,在这两类受试者中,刚才描述的低估效应与高估效应是

存在的。第二,玩家人数的增多,因而他们中的竞争的增加,往往增加了对高概率的低估和低概率的高估,这是一个有点令人惊奇的结果。

247 与赛事结束之后后验地决定的获胜的客观概率相比较,按照同注分彩制(parimutuel system)赌马提供了主观概率和客观概率的现成的、真实生活中的比较。在一场赛事中,为一匹马下的总赌注除以为所有的马下的有效的总赌注,决定了成功的可能性,比如,这匹马获胜的集体的主观概率(the collective subjective probability)。格里菲斯(Griffith,1949)在1947年考察了1 386场赛事中的数据;他的结果与普雷斯顿和巴拉塔所获得的那些结果在定性的意义上明显一致。这些主观概率系统地高估了低的客观概率,系统地低估了高的客观概率。就格里菲斯的数据而言,客观概率和主观概率在0.16是相等的,这接近于普雷斯顿和巴拉塔获得的0.20的值。格里菲斯在最后的脚注中评论说,对1934年8月在相同赛马跑道上举行的所有赛事进行相同的计算,基本上获得了相同的结果。在这种情况下,等价点落在0.18不是0.16。从1934年的经济萧条到1947年的经济相对繁荣,这些结果的不变性,增加了它们的重要性。

还有一个与格里菲斯非常相似的更加详尽的研究,是由麦格洛斯林(McGlothlin,1956)对从1947年到1953年大多数在加利福尼亚赛马场举行的9 605场赛事的计算提供的。正如人们将期望的那样,他的结果的一般特征与较早研究的那些结果是接近一致的。客观概率和主观概率曲线相交在0.15和0.22之间。麦格洛斯林在特定一天的通常的8场赛事中也发现了一些有趣的趋势。首先,高估低概率,有增加的趋势,而且,这种现象在当天的决赛中是特别令人吃惊的。在当天的决赛中,许多玩家似乎明显地对不足以挽回前面赛事造成的损失的可能性不

感兴趣。这种现象是如此令人吃惊,以至于即使在减去场地费(当时在加利福尼亚是 13%)之后,为低可能性的马下赌注的人,也有一个有效的正的期望值。

户田(Toda,1951,1958)提出的测度主观概率的双人游戏方法与普雷斯顿和巴拉塔的拍卖程序非常相似。为了获得对基本事件和复合事件的主观概率的扩展测度,舒福德(Shuford,1959)提供了对这种方法的更加有趣的应用之一。受试者把一枚二十面的骰子(表面上的整数是从 1 到 20)摇两次,选择一个纵轴和横轴的 20×20 矩阵的行与列。受试者是在空军基地接受过训练的六十四名飞行员,他们大体是成对的。在一个特定试验中的事件序列如下。把这个横轴和纵轴的矩阵投影在一个屏幕上。受试者 A 写下他的叫价 $x(0\leqslant x\leqslant 10)$,比如说,他选择横轴;受试者 B 决定"买"或"卖"A 的叫价;实验者通过掷骰子决定赌注,即选择投影矩阵的哪个元素;最后,受试者记下他们在游戏中的得分。能够表明,A 在这场游戏中的乐观(最低限度的)策略是使 x 等于 10 乘以他对有利结果的主观概率。

每一对受试者都玩两场游戏。在第一场游戏中,当通过把这个二十面的骰子摇两次来选择这个矩阵的位置时,结局是由根据情况所出现的横轴或纵轴来决定的。在第二场游戏中,如果对矩阵元素的两次连续选择,是在两个同类轴上掷骰子的结果,打赌就会获胜。

绝大多数受试者高估低概率和低估高概率,证实了普雷斯顿和巴拉塔、格里菲斯以及麦格洛斯林较早的发现。对于基本事件和复合事件来说,都是如此。另一方面,舒福德发现,许多受试者对主观概率的估计,很符合客观概率的一个线性函数,尽管这个函数的斜率和截距随着受试者的变化而变化。他发现的对复合事件的估计,是特别有趣的。绝大多数的受试者接近正 248

确的规则，即他们把复合事件的概率估计为近似于基本事件概率的平方。对这种正确规则的应用是如此共同的，这一点是令人吃惊的，因为在一系列试验结束后，当问这些受试者，他们用了什么规则时，只有两个人指出了正确的规则。

从所引证的研究成果来看，显然，在经验上测度主观概率是有可能的，此外，这样做的显而易见的方法是，根据个人做出的打赌类型。另一方面，回想一下，上面提供了对这种测度方法的反对意见，因为当赌额发生变化时，打赌的可能性也是变化不定的。他们根据预期效用概念，促进了更深入的理论发展，我这里将不研究这个问题。参考文献可参见，Savage(1954)、Luce and Suppes(1965)、Fishburn(1970)、Krantz et al.(1971，第 8 章)以及 Luce(2000)。下面的一小节讨论萨维奇的公理，但不是他的基本表征定理。

信念的不准确衡量：较高概率和较低概率。[①] 在主观概率传统或贝叶斯统计程序中思考衡量信念问题的几乎每一个人都承认对下列问题感到不安：即总是在对概率的先验估计中要求精度到下一位小数，或者，要求决定事件概率分布的参数。另一方面，由拉姆齐(Ramsey，1931)、德·菲内蒂(1931，1937/1964)、库普曼(1940a，b)、萨维奇(1954)和后来的作者提出的在不确定条件下理性地作出决策的形式理论，几乎一律倾向于产生出一种结果，这种结果确保用自然状态或实体的任何其他集合的惟一概率分布来表达先验的信念。我考察了这些标准理论中的一些理论，而且提出了我们如何能够更好地批评他们作出的断言的问题。此外，我考虑了这样的主张：能够把这些公理中表达的理想化看成是纯理性的理论。

---

① 这一小节的大部分内容取自苏佩斯(1974d)。

因为前面提到的标准理论基本上得出了相同的形式结果，也就是说，存在着关于自然状态的惟一的概率分布，所以，对一个理论的批评几乎完全适用于对全部理论的批评。由于这种理由，我们可以集中关注于萨维奇的(1954)公理，因为很多读者熟悉这些公理，也因为这些公理已经在文献中进行过许多讨论。然而，我强调，我对萨维奇的公理所不得不说的东西在不作改变的情况下本质上也适用于其他标准理论。

因为从形式的观点来看，萨维奇的公理是相当复杂的，所以，我在这里将不对它们作出明确的阐述，但将试图描述它们的直观内容。这些公理是关于决策中的偏好的，这种情况下，决策是从自然状态的集合到结果集合的映射或函数。让我用我过去举过的一个事例来说明这些思想(Suppes 1956)。 249

某位独立的面包经销商必须在前一天晚上的十点之前下订单。他的销售与食品商无关，受到交货时天气是否下雨的影响，因为如果天气下雨，食品商往往根据他们有较少顾客的累积凭证进货较少。在下雨天，经销商最多能卖出七百块面包；在这一天，如果他预订的面包多于七百块，他就少赚钱。另一方面，当天气晴朗时，他能够卖出大约九百块面包。如果作出这种简单化的假设：他可能只是根据他的净利润总结出用一种已知的自然状态(s_1 -下雨或 s_2 -不下雨)作出已知决定的结果，表 1 表示他面临的情况。经销商的问题是做出一个决策的问题。

表 1　买进面包的决策矩阵

	d_1	d_2	d_3
	买 700 块	买 800 块	买 900 块
s_1 -下雨	\$21.00	\$19.00	\$17.00
s_2 -不下雨	\$21.00	\$24.00	\$26.00

显然,如果他肯定知道,明天将下雨,他应该作出决策 d_1,如果他肯定知道,明天不下雨,他应该作出决定 d_3。把萨维奇的理论扩展到更一般和更复杂情境的问题是,确信这些公理的任何一个人都将使期望效用最大化,以这样一种方式把公理强加于这些决策中的选择或偏爱。这意味着,他相信这些公理的方式将会产生他关于自然的真实状态的信念的一种主观概率分布,以及结果集合上的一个效用函数,以致相对于关于自然状态的一种主观概率分布和结果集合上的一个效用函数,以一种直接的方式,定义一个已知决策的期望值。正如人们预料的那样,萨维奇事实上在他的第一个公理中要求,在决策中的偏爱是可传递的,而且,已知任何两种决策,至少偏爱一种决策的程度比偏爱另一种决策的程度要弱。当决策的定义域限于一个已知的自然状态的集合时,公理 2 把这种排序假设延伸到保持的相同特性;例如,决策者可能知道,自然的真实状态在于全集的某个子集。公理 3 断言,在根据决策中的偏爱来定义结果中的偏爱的地方,一个事件的知识不可能改变结果中的偏爱。公理 4 要求,已知任意两个自然状态的集合,即任意两个事件,一个至少与另一个一样可能,也就是说,在事件中的定性概率是强关联的。公理 5 排除了下列无价值的情况:在这些情况下,所有的结果在效用上都是等价的,因此,每一个决策都等价于另一个决策。公理 6 基本上是说,如果事件 A 不如事件 B 有可能(A 和 250 B 是相同自然状态集的子集),那么就存在着对自然状态的一种划分,使得根据 A 划分的每个元素的并集,不如根据 B 划分的每个元素的并集有可能。众所周知,萨维奇的这个公理是与德·菲内蒂和库普曼的公理密切相关的,德·菲内蒂和库普曼的公理要求,把现有的自然状态划分成任意多的等概率事件。最后,他的最后一个公理,即公理 7,阐述了确实原则(sure-

thing principle)。

我的第一个主要主张是,萨维奇的有些公理,在任何一种直接的意义上,都不代表任何一位理想理性的人都应该满足的理性公理,而是相反,它们代表了可以满足也可以不满足的关于环境的结构假设。

许多年前,在伯克利的第三次专题讨论会上(1955),我介绍了决策理论中的结构公理和理性公理之间的区别(Suppes,1955)。从直观意义上看,与理性公理相对的结构公理,在特征上,是关于存在的。在萨维奇的七个公设的情形中,有两个公设(5 和 6)是结构公理,因为它们在特征上是关于存在的。

如果有一枚硬币,一位决策者相信,对于任何一个有穷投掷序列,这枚硬币都是公平的,萨维奇通过坚持它是可应用的,为他的强公理 6 辩护。然而,对这种论证有几种反对意见。首先,如果认真地采纳这种论证,那么人们应该重新整理整个基础,只围绕 $p=0.5$ 的伯努利序列来创建这个有穷投掷序列,并达到任意地逼近任何一个预期事件的概率。更重要的是,在根本不改变人类思维的条件下,站在人类的立场上,在评价他们关心的有意义的事件的可能性或概率,不管是定性的或定量的,思考掷一枚硬币的有穷序列,简直是不正常的。

考虑这样一个例子,一位病人决定,是否遵从外科医生的劝告,做重要的外科手术。让我们假设,外科医生已经估计到手术的有利方面和不利方面,而且,现在这位病人面临着做出关键的决定:是冒着至少有死亡的正概率的风险做手术呢,还是冒着不做手术的风险,继续承受患病的后果呢。我发现,站在病人的立场上,认为思考掷一枚公平硬币的有穷序列,对于做出理性的决定将有任何帮助是完全不可能的,并且,在心理学上也是很不现实的。

另一方面,其他公理,像关于偏爱的排序或定性概率的那些公理一样,在这个框架内似乎是合理的,也是不难接受的。但这是重要的观点。在不确定性起着核心作用的情况下,在实践中做出决定,没有任何企图要达到这样的状态:拥有一个替代的定量概率估计,或者,如果你喜欢的话,拥有一个可计算的期望效用。

事实上,我坚信的是,我们通常研究受限制的情况,在这些情况下,供我们选择的决策集合很小,以及在这些情况下,我们认为相关的事件在数量上并不多。标准理论提供的这种扩大了的决策框架,恰好是本小节的第一句话中提到的那种不安的来源。从直观上看,我们所有的人都改变了任意精致地估计概率的想法。为了拥有一个足够大的决策空间,我们也改变了引入一个详尽的随机化机制的想法。的确,已知贝叶斯关于随机化的态度,采用萨维奇的掷一枚公平硬币的有穷序列方式,造成了一种悖论的氛围。

251

似乎对我而言,表明这个问题的另一种方式是,存在着一个很强的直观感:一位决定者不是非理性的,只是因为范围广泛的决策可能或事件对他来说并不存在。要求决策者扩大他的决定空间,例如,增加一枚可被投掷任意有限次数的硬币,并不是理性的一个组成部分。我感到,固有的理性理论应该是准备研究一个已知的自然状态集和已知的决策函数集,而且,信念或决策的形式理论的责任是提供这样一个理论:如何不引入强结构假设来研究这些受限制的情况。

表达我所说的关于纯理性公理这个短语的一种技术方式如下。为了保持技术方法的简单性,暂时,让我们限于自然状态的基本集合 S 和在 S 的子集上的定性概率的一种二元排序关系,包括从事件角度看具有其直观意义的通常的并集、交集和补集

的布尔运算。于是,我说,关于这些结构的一个公理,只有在子模型的条件下是闭合的,就是一个纯理性的公理。从技术上看,在子模型下的闭合意味着,如果一个对(S, $\geq$)满足这个公理,那么,具有二元关系$\geq$ 的 S 的任何一个非空子集,都满足这个公理,这种二元关系限于这个子集的幂集,即限于已知子集的所有子集集合。(当然,并集、交集和补集运算,在这个子集的幂集中,是闭合的。)运用这种技术定义,我们能很容易看到,在萨维奇的七个公理的定义中,有五个定义满足这种限制,而已经作为结构公理提到的两个定义,不满足这种限制。

让我试着使支持下列必要条件的直觉变得稍微更明显一些:这个必要条件是,纯理性的公理应该满足子模型下的闭合条件。这种条件的一种应用接近于在选择理论中有关不相关替代项间相互独立性的公理。这个公理说,比如,如果我们在所提供的候选人中间表示偏爱,而且,如果一位候选人由于死亡或其他原因被从名单上除名,那么我们在其余的候选人中间的偏爱排序应该是不变的。这个公理满足子模型下的闭合条件。核心思想是,追求和作出关于环境的特殊要求的存在判断的必要条件,不代表纯理性的要求,而是代表关于环境的结构要求,而且,子模型下的闭合条件排除了这样的存在判断的要求。

一个不同的但密切相关的定义纯理性公理的方式是,这样一个公理一定是预期的数值表征的存在性的一个逻辑结果。这个标准,我们能够称之为**表征结果的标准**,能够被接受为是必要的和充分的,而在子模型下闭合的标准显然不是充分的。另一方面,表征结果的标准的外在特征,可能被认为是不能令人满意的。有用的是识别这样的公理:这些公理对于这种预期表征来说是不必要的,因此,是私自引入了某种不想要的任意结构假设。应该明白,萨维奇的公理 5 和 6 就是这样私自引入的。

252 我很愿意承认这种观点：我们能够考虑更受限制的那种理性公理。人们可能认为，我们针对特殊的情况，需要特殊的理性公理，而且，我们应该以下列情况的一种分类法为出发点：这些情况为每一种主要类型的分类法提供了适当的公理。然而，在当前的主要分析状态中，似乎令人向往的是，以明显地区分理性与结构公理为出发点，并拥有特征上普遍的纯理性的概念。

现在，返回到我对萨维奇理论的批评，很容易为萨维奇的五个理性公理提供有穷或无穷模型，对于这五个理性公理来说，根据效用和主观概率，根本不存在数值表征。用我这里所用的语言，萨维奇的纯理性公理不足以确立：存在着表征的数值效用和主观概率函数。

此外，我们可以表明，一个普遍特征的附加基本公理的有穷名单，没有一个足以保证存在着适当的数值函数。我所说的**基本公理**是指，能够在一阶逻辑的范围内来表示的公理。正如第2章所说明的那样，一阶逻辑基本上是由语句连词、同一层次的变量和这些变量的量词，以及非逻辑谓词、运算符号和个别常量的概念框架构成的。例如，群或有序的代数域的标准公理是基本的，但实数域的最低上限公理不是基本的。以一种基本的方式阐述萨维奇的公理5，而不是公理6，是有可能的。

在无穷模型的情况下，基本公理的不充分性，如果没有限制它们是一个普遍特征的话，是从上升的勒文海姆—斯科尔姆—塔斯基定理(Löwenheim-Skolem-Tarski theorem)再加上某些弱的一般假设得出的结论。这个定理断言，如果一个基本公理集拥有一个无穷模型(即其定义域是一个无穷集的模型，萨维奇的理论正是这种情况)，那么它拥有每一个无穷大基数的一个模型。在很一般假设的条件下，例如，关于偏爱或更大的主观概率的排序关系，不可能把高次无穷大基数的模型映射到实数中，因

而根本不存在数值表征。

在有穷模型的情况下,2.2 节讨论的斯科特和苏佩斯(1958)的方法适用于表明,普遍的基本公理的有穷集没有一个满足要求。由萨维奇的五个纯理性公理组成的体系拥有有穷模型,但根据已经指出的方法,我们能够表明,用强的足以导致标准数值表征的理性的普遍公理,根本没有有穷的基本扩展。

上限概率和下限概率的定性公理。 我完全根据我考虑的信念或主观概率,引入对萨维奇公理的适当的有穷论的类比。这些构成了对定义 1 中陈述的德·菲内蒂的定性公理的一种扩展,并导致了对任意事件中的信念的简单近似的衡量。这些公理要求我前面部分地批评过的那种观点,即存在着其概率是完全已知的事件的某个标准集合。例如,已知某个确定的 n,把一枚公平的硬币掷 n 次,满足这些公理。它们不要求 n 是无穷大, 253 因此,可以把 n 看成是更现实的。我这里给出这些公理,尽管我的感觉是,像前面提到的是否做外科手术的决定一样,从作出严肃的决策的观点来看,它们可能是不令人满意的。但它们确实提供了德·菲内蒂的思想与关于无穷划分的标准结构公理的有穷看法的一种结合。

上限概率和下限概率(upper and lower probabilities)的概念在文献中似乎是相当新近的,但它显然与由卡拉特奥多里(Caratheodory)等人在 19 世纪末和 20 世纪初提出的内测度与外测度的经典概念密切相关。库普曼(1940b)明确地引入上限概率和下限概率,但从概念的观点来看,却对上限概率和下限概率没作考察。正如在数学分析中的上限测度和下限测度的情况下那样,他把上限概率和下限概率用作定义概率的一个技术方法。第一个明显的概念讨论似乎是相当新近的(Smith, 1961; Good, 1962)。史密斯(Smith)专门研讨了许多重要的概念考

虑,而且,古德陈述了似乎是自然强加于上限概率和下限概率的许多定量特性。应用于统计推理的问题,可在登普斯特(Dempster)的文章(1967)中找到。然而,就我所知,在文献中确实还没有从完全定性的公理出发的其他简单的公理化处理,下面给出的这些公理代表了这样一种努力。它们显然不可能是最一般的公理,但它们确实提供了一个简单而相当优雅的定性基础。

从形式的观点来看,这些公理应用的基本结构是四元组$(\Omega, \mathfrak{I}, \mathcal{S}, \succeq)$,其中,$\Omega$是一个非空集,$\mathfrak{I}$是$\Omega$子集的一个代数,也就是说,$\mathfrak{I}$是$\Omega$的一个非空子集族,而且,在并集与补集下是闭合的,$\mathcal{S}$是一个相似的集合代数,直观上用来表示标准测度的事件,我把$\mathcal{S}$中的事件看成是标准事件S、T,等等。关系$\succeq$是在$\mathfrak{I}$上的熟悉的排序关系。我用熟悉的缩写表示等价和根据弱排序关系的严格排序。(正如前面提到的,一种弱排序是可传递的和强关联的,即对于任何事件A和B,或者$A \succeq B$,或者$B \succeq A$。)

定义 4. 一个结构$\mathbf{\Omega} = (\Omega, \mathfrak{I}, \mathcal{S}, \succeq)$是对信念的一个有穷的近似的衡量结构,当且仅当,$\Omega$是一个非空集,$\mathfrak{I}$和$\mathcal{S}$是在$\Omega$上的集合代数,而且,对于$\mathfrak{I}$中的每个$A$、$B$和$C$和$\mathcal{S}$中的每个$S$和$T$来说,都满足下列公理:

公理 1. 关系$\succeq$是$\mathfrak{I}$的一种弱排序;

公理 2. 如果$A \cap C = \varnothing$和$B \cap C = \varnothing$,那么$A \succeq B$,当且仅当,$A \cup C \succeq B \cup C$;

公理 3. $A \succeq \varnothing$;

公理 4. $\Omega \succ \varnothing$;

公理 5. $\mathcal{S}$是$\mathfrak{I}$的一个有穷子集;

公理 6. 如果$S \neq \varnothing$,那么$S \succ \varnothing$;

公理 7. 如果 $S \succeq T$，那么在$\mathcal{S}$中存在着一个 V，使得 $S \approx T \cup V$。

在比较公理 3 和 6 时，注意，A 是一般代数 $\mathfrak{F}$ 的一个任意元素，而事件 S(公理 6 中提到的)是子代数$\mathcal{S}$的一个任意元素。同样，在公理 7 中，S 和 T 是子代数$\mathcal{S}$中的标准事件，不是一般代数中的任意事件。公理 1—4 恰好是熟悉的没有任何改变的德·菲内蒂公理。因为所有的标准事件(在数目上是有穷的)也是事件(公理 5)，所以，公理 1—4 对于标准事件和任意事件都成立。公理 6 确保子代数$\mathcal{S}$的每个最小的元素都有正的定性概率。在技术上，$\mathcal{S}$的一个最小元素是$\mathcal{S}$中的任意事件，使得 $A \neq \varnothing$，但不是这种情况：在$\mathcal{S}$中存在着一个非空的 B，使得 B 是 A 的一个真子集。$\mathcal{S}$的一个最小开区间(S, S')是使得 $S \prec S'$ 和 $S' - S$ 等价于$\mathcal{S}$的一个最小元素。公理 7 是主要结构公理，它对于子代数成立，而对于一般代数不成立；它为标准事件阐述了一个极其简单的可解性条件。苏佩斯在较早的一般代数的情况下以这种形式陈述过公理 7(1969a, p. 6)。 254

在陈述满足定义 4 的结构表征定理和惟一性定理时，除了关于标准事件的普通概率测度之外，我将用上限概率和下限概率来表示对任意事件的不精确的测度。对人们预期的这种上限概率和下限概率的定量特性的一个好的讨论，可以在古德的文章(1962)中找到。他提出的所有特性在这里都是不需要的，因为他处理条件概率。下列特性是基础的，其中，(对于 $\mathfrak{F}$ 中的每个事件 A 和 B 来说)$P_*(A)$是一个事件 A 的下限概率，$P^*(A)$是上限概率：

Ⅰ. $P_*(A) \geqslant 0$。

Ⅱ. $P_*(\Omega) = P^*(\Omega) = 1$。

Ⅲ. 如果 $A \cap B = \varnothing$，那么 $P_*(A) + P_*(B) \leqslant P_*(A \cup$

$B) \leqslant P_*(A) + P^*(B) \leqslant P^*(A \cup B) \leqslant P^*(A) + P^*(B)$。条件(Ⅰ)对应于古德的公理 D2,条件(Ⅲ)对应于他的公理 D3。

对于标准事件,$P(S) = P_*(S) = P^*(S)$。对于在定性概率方面不等价于一个标准事件的一个任意事件 A 来说,我把它的“真”概率看成是位于开区间$(P_*(A),\ P^*(A))$中。

最初,作为第四个特性,我包括

$$P_*(A) + P^*(\neg A) = 1$$

其中,$\neg A$ 是 A 的补集,但扎诺蒂(Mario Zanotti)对我指出,这个特性是通过下列论证从(Ⅱ)和(Ⅲ)推出的:

$$\begin{aligned} 1 = P_*(\Omega) &= P_*(A \cup \neg A) \\ &\leqslant P_*(A) + P^*(\neg A) \\ &\leqslant P^*(A \cup \neg A) \\ &= P^*(\Omega) = 1 \end{aligned}$$

在定理 7 的第四部分,我定义一个确定的关系,并且主张,它是一个具有源于对上限概率和下限概率的不等式成立的半序关系的蕴含的半序。半序作为对简单序的一种概括在文献中已经得到了相当广泛的讨论,第一个讨论是由卢斯(Duncan Luce)提供的。我这里运用了由斯科特和苏佩斯(1958)给出的公理。一种结构$(A,\ *\succ)$(其中,A 是一个非空集,$*\succ$是在 A 上的一种二元关系)是一种半序,当且仅当,对于所有的 $a,\ b,\ c,\ d \in A$:

公理 1. 非 $a * \succ a$;

公理 2. 如果 $a * \succ b$ 和 $c * \succ d$,那么,要么 $a * \succ d$,要么 $c * \succ b$;

255 公理 3. 如果 $a * \succ b$ 和 $b * \succ c$,那么,要么 $a * \succ d$,要么

$d * > c$。①

定理 7. 设 $\boldsymbol{\Omega} = (\Omega, \mathfrak{J}, \mathcal{S}, \succeq)$ 是信念的一个有穷的近似的衡量结构，于是

(i) 存在在$\mathcal{S}$上的一个概率测度 P，使得对于任何两个标准事件 S 和 T，

$$S \succeq T, \text{当且仅当}, P(S) \geqslant P(T)$$

(ii) 测度 P 是惟一的，并且把相同的正概率赋予$\mathcal{S}$的每个最小事件。

(iii) 如果我们把 P_* 和 P^* 定义如下：

(a) 对于等价于某个标准事件 S 的 $\mathfrak{J}$ 中的任何一个事件 A，

$$P_*(A) = P^*(A) = P(S)$$

(b) 对于不等价于某个标准事件 S，但位于标准事件 S 和 S'最小开区间(S, S')的 $\mathfrak{J}$ 中的任何一个事件 A，

$$P_*(A) = P(S) \text{ 和 } P^*(A) = P(S')$$

于是，P_* 和 P^* 满足在 $\mathfrak{J}$ 上的上限概率和下限概率的条件(Ⅰ)—(Ⅲ)，并且

(c) 如果 n 是在$\mathcal{S}$中的最小元素的个数，那么对于 $\mathfrak{J}$ 中的每个 A，

$$P^*(A) - P_*(A) \leqslant 1/n$$

---

① 在所引证的这篇文章中，已经表明，当 A 是一个有穷集时，那么能够为半序证明下列表征定理。存在着在 A 上定义的一个数值函数，使得对于 A 中的所有 a 和 b，

$$\varphi(a) > \varphi(b) + 1, \text{当且仅当}, a * > b$$

然而，如果 A 是一个无穷集，这种表征并不一定成立，因此，不能把它用在定理 7 下面的条款(iv)中。

(iv) 如果我们在 $\mathfrak{I}$ 上定义 A 和 B,

$A * \succ B$,当且仅当,在$\mathcal{S}$中 $\exists S$,使得 $A \succ S \succ B$,

那么 $* \succ$ 是在 $\mathfrak{I}$ 上的一个半序,如果 $A * \succ B$,那么 $P_*(A) \geqslant P^*(B)$,而且,如果 $P_*(A) \geqslant P^*(B)$,那么 $A \succeq B$。

证明: 一旦观察到,子代数$\mathcal{S}$同构于一个有穷的集合代数,它含有对应于单位集合(即原子事件)$\mathcal{S}$的最小事件,部分(i)和(ii)就是从在苏佩斯(1969a, pp. 7 - 8)的文章中给出的证明得出的。

至于部分(iii),直接证实了上限概率和下限概率的条件(Ⅰ)和(Ⅱ)。为了证实条件(Ⅲ),假设 A 和 B 都不等价于一个标准事件,就足够了,因为如果两者之中有一个等价于一个标准事件,就简化了这里给出的论证,如果两者都等价于一个标准事件,就立即从标准测度 P 的特性中推出(Ⅲ)。因此,我们可以假设,A 位于一个最小区间(S, S')中,B 位于一个最小区间(T, T')中,即 $S \prec A \prec S'$ 和 $T \prec B \prec T'$。既然根据(Ⅲ)的假设,$A \cap B = \varnothing$,所以 $T \preceq \neg S$,因为如果 $T \succ \neg S$,我们将有 $A \cup B \succ S \cup \neg S$,这是不可能的。现在,很容易检查,对于标准事件,如果 $T \preceq \neg S$,那么,在$\mathcal{S}$中 $\exists T^*$,使得 $T^* \approx T$ 和 $T^* \subseteq \neg S$。因此,我们有

$$P_*(A) + P_*(B) \leqslant P(S) + P(T^*) = P(S \cup T^*)$$

$$\leqslant P_*(A \cup B)$$

含有从 $S \cup T^* \prec A \cup B$ 推出的最后的不等式,$S \cup T^* \prec A \cup B$ 本身是$S \prec A$、$T^* \prec B$, $A \cap B = \varnothing$ 和公理 2 的一个直接结果。对于下一步,如果在$\mathcal{S}$中 $\exists T^{**}$,使得 $T^{**} \approx T'$ 和 256 $T^{**} \subseteq \neg S'$,那么$A \cup B \prec S' \cup T^{**}$,而且,设$A \cup B$是在最小的闭区间$[V, V']$中,即$V \preceq A \cup B \preceq V'$。于是,很容易表明,

$V \prec S \cup T^{**}$，因而

$$P_*(A \cup B) = P(V) \leqslant P(S \cup T^{**})$$

$$= P(S) + P(T^{**})$$

$$= P_*(A) + P^*(B)$$

并且，既然 $S \cup T^* \prec A \cup B$ 和 $V \preceq S \cup T^{**}$，所以，要么 $A \cup B \preceq S \cup T^{**}$，要么 $A \cup B \preceq S' \cup T^{**}$，在任何一种情况下，

$$P_*(A) + P^*(B) = P(S \cup T^{**}) \leqslant P^*(A \cup B)$$

$$\leqslant P(S' \cup T^{**})$$

$$= P(S') + P(T^{**})$$

$$= P^*(A) + P^*(B)$$

另一方面，如果不存在 T^{**}，使得 $T^{**} \approx T'$ 和 $T^{**} \subseteq \neg S'$，那么 $T' \succ \neg S'$，因此，$S \cup T^* = S \cup \neg S$，结果，$A \cup B \approx S \cup T^*$，所以，$A \succeq S$ 或 $B \succeq T^*$ 与假设相反，这就完成了(Ⅲ)的证明。

部分(iii)的(c)的证明立即从(ii)和(iii)的前面部分推出。(iv)的证明也是简单明了的，将被省略。

我在开场白中提到的一个令人为难的问题是：质问一个主观概率的下一位小数点的问题。没有主张应对所有这些问题，定理 7 体现的结果表明，定义 4 的公理提供了一个更好回答的基础。如果存在着 n 个最小的标准事件，那么 2^n 个标准事件的概率恰好被认为是形式 m/n（其中，$0 \leqslant m \leqslant n$）的有理数，而且，关于精确性的进一步质问是错误的。所有其他事件的上限概率和下限概率都是根据 2^n 个标准事件的概率来定义的，因

此，上限概率和下限概率也恰好被认为是相同形式 m/n 的有理数。

最后，我明确地指出，在定义 4 中没有必要要求，样本空间 Ω 是有穷的。惟一基本的必要条件是，标准事件的集合 $\mathcal{S}$ 是有穷的。代数 $\mathfrak{I}$ 甚至能够拥有大于连续统的基数的基数，因此，在 $\mathfrak{I}$ 上的排序关系 $\succeq$ 可能在数值上不是可表征的，而且，在 $\mathfrak{I}$ 中的所有事件的上限概率和下限概率，像在这个定理中一样，是存在的和明确的。

5.8 结语：关于概率的实用主义

对表征概率的主要概念方式的这么长的概述，可能会令人误导地建议说，当科学家们运用概率概念时，他们选择自己喜欢的表征作为基础。就像在具有类似特征的其他基本问题的大多数情形中一样，科学实践是相当不同的。也许，我们最自然地指望为某个概率表征找到一种基本承诺的两门学科是统计力学和量子力学，但情况完全不是这么一回事。我将不试图详细地为我的想法提供引证，自麦克斯韦以来的文献对统计力学来说表明了什么，或者，自海森堡和狄拉克以来的文献对量子力学揭示了什么。我将仅给出一些举例说明。

早期的统计力学。在早期岁月里，比如说，在麦克斯韦于 1860 年和 1867 年关于气体动力学理论的著名文章中，存在着各种不同的概率计算，但根本没有提到应该如何解释概率的问题。从 1860 年到 20 世纪末①，关于如何能够在统计物理学中

257

① 原文是"the end of this century"，为了明确，这里译为 20 世纪末。——译者

运用概率概念的问题,演化出一种更精致的分析。像玻尔兹曼(Boltzmann)、吉布斯、洛施密特(Loschmidt)、基尔霍夫(Kirchhoff)、普朗克(Planck)、彭加勒和策梅洛(Zermelo)之类的著名人物,都在很大程度上详尽地论证过统计物理学的基础,但不是关于概率的基础。对这些问题的一个好的讨论和广泛的参考文献可以在布拉什(Brush)的著作(1976)中找到。

量子力学。在关于量子力学的早期论文中,对待概率的一种更严峻的态度占有统治地位。海森堡 1925 年的著名论文和狄拉克于 1926 年初发表的两篇文章根本没有提到概率。[再版的量子力学的这些论文和其他早期论文可以在范·德·韦尔登(van der Waerden)主编的论文集(1968)中找到。]狄拉克在他的关于量子力学的著名论述中(1949,第 47—48 页)明确地讨论了,如何计算可观察量的平均值和有特定值的一个可观察量的概率问题,但只字未提如何解释概率的问题。

福克。一个哲学上更加深刻的和更令人感兴趣的讨论,能在著名的俄罗斯量子物理学家福克(V. A. Fock)的著作(1931/1978)中找到,但他的分析不符合本章前面讨论的概率的各种主要解释。他认真地强调了两点。第一点是,量子力学中的概率是对潜在可能性的一种表达,第二点是,如果我们完全承认电子的波粒二象性,理解海森堡不确定关系的概率困难就会消失(第 93 页)。这个简要的总结没有对福克的观点给出一种丰富的或更公平的概述,因此,让我提供一些详尽的引文,这些引文是某些科学哲学家不熟悉的。这里是关于潜在可能性的很好的一段:

> 如果我们把原子客体与测量仪器之间的相互作用的行为看成是我们判断客体特性的来源,而且,如果在研究现象时,我们考虑到相对于观察手段的相对性概念,我们就在本

质上把一个新的元素引入到对原子客体及其态与行为的描述中，即概率的思想，因此是潜在可能性的思想。把概率概念看成是描述的重要元素，而不是我们知识的不完备的一个信号，这种需要是从下列事实中得出的：对于已知的外部条件，客体与仪器的相互作用的结果，一般说来，不是被惟一地预先确定的，而只有一个确定的发生概率。根据客体的固定的初始态和已知的外部条件，一系列这些相互作用导致了对应于一种确定的概率分布的统计数据。这个概率分布反映了在已知条件中存在的潜在可能性。

(Fock, 1931/1978: p. 19)

下面是关于统计系综和不确定关系的很好的一段：

在量子力学发展的初期，物理学在寻找统计(概率)解释的早期努力中，仍然受到了电子是一个经典粒子的概念的束缚。甚至当德布罗意提出物质的波动性思想时，物质波有时被解释为携带粒子的某种东西。后来，当海森堡的关系出现时，它们被解释为是**不精确的关系**，而不是不确定关系。比如说，物理学家认为，电子有确定的位置和速度，但根本没有可能确定其中的任何一个。波函数的模的平方被解释为具有已知坐标的(坐标被认为是明确的)一个粒子的概率密度——与实际的实验条件无关。对动量空间中的波函数的模的平方也给出了相似的解释。(在普通空间与动量空间中)这两种概率都被同时认为是一个确定的复合事件的概率，特别是粒子具有明确的坐标值与动量值。因此，通过海森堡关系表示的同时测量两个值的实际上的不可能性，体现为一个悖论或大自然的反复无常的变化，据此，并非存在的一切都是可认识的。

258

如果我们完全承认，电子的波粒二象性确定了它的本质，并掌握了量子力学概率指什么和它们属于怎样的统计系综，那么所有这些困难就会消失。

首先，让我们试图给出一个统计系综的一般定义。我们假设具有不同特征的元素的一个无限制的集合，这使得有可能对这些元素进行分类，并观察具有已知特征的一个元素出现的频率。如果就此而言，存在着一个明确的概率（即对于这个集合的每个元素来说），那么这个集合就构成了一个统计系综。

同在经典物理学中那样，在量子力学中，惟一能够考虑的集合是其元素具有明确的参量（特征）值的那些集合，据此，能够进行分类。这意味着，一个统计系综的元素必须是用经典语言描述的，而且，一个量子客体不可能是一个统计系综的一个元素，即使能够把波函数赋予这个客体。

在量子力学中考虑的统计系综的元素不是微观客体本身，而是对它们的实验的结果，一个明确的实验安排对应于一个明确的系综。这些结果被在经典意义上加以描述，因此，能够作为分类这个系综元素的基础。既然对于不同的量来说，从一个已知波函数产生的概率分布对应于不同的实验安排，所以，它们属于不同的系综。

(Fock，1931/1978：pp. 93 - 94)

稍后福克关于相对频率确实是这么说的。

具有已知初始状态的一个客体的这种或那种行为的概率，是由该客体的内在特征和外在条件的本质所决定的；它是描写该客体的这种或那种行为的潜在可能性的一个数

目。于是,这个概率本身表现出该客体的一种已知行为发生的频率;相对频率是它的数值测量。因此,这个概率在本质上属于单个客体(而不是客体的系综),并且描写了它的潜在可能性。同时,人们为了根据实验确定它的数值,必须有实现这些可能性的统计,所以,实验必须被重复多次。由此可知,量子理论的概率特征没有排除这个事实:量子理论是以单个客体的特性为基础的。

259　　总结起来,我们能说量子力学的主要概念(即由波函数描述的态的概念)的目的,就是客观地分析微观客体固有的所有潜在可能性。这决定了这个理论的概率本性。

(Fock, 1931/1978: pp. 94 - 95)

在最后这一段,福克顺便提到了相对频率,他对这个概念的处理方式是标准的和浮浅的,与他对其他问题的探索性评论相反。相比之下,让我们看一下两位杰出的数学家的陈述,他们在量子力学基础方面的工作是很著名的。

韦尔。首先,韦尔(Weyl, 1928/1931)对相对频率的论述如下。这应该对所有实验科学都成立,不只是对量子力学乃至物理学都成立。

对于实验科学来说,概率的意义是,它们决定在一系列**重复的观察中出现的相对频率**。按照经典物理学,原则上,有可能创造一些条件,在这些条件下,与一个已知的物理系统相联系的每个量假设一个任意明确确定的值,只要这些条件是相同的,这个值就恰好是可重现的。**量子物理学否定这种可能性。**

(Weyl, 1928/1931: p. 75)

韦尔继续在下一页明确地用了康德关于自然科学的评论。

> 自然科学具有构造的特征。它研究的概念不是通过直接认知能够从客观世界获得的那些性质或属性。这些性质或属性只能是通过一个间接的方法论,通过观察它们对其他物体的作用来决定的,相应地,它们隐含的定义是,以支配作用的明确的自然定律为条件。例如,考虑伽利略的质量概念的引入,它基本上相当于下列间接的定义:"每个物体都有动量,即具有与它的速度 v 的方向相同的一个矢量 mv;标量因子 m 被称为它的质量。一个封闭系统的动量是守恒的,也就是说,许多作用物体的动量之和,在作用前后是相同的。"当把这个定律应用于被观察的碰撞现象时,数据是可得到的,它们允许确定不同物体的相对质量。但是,科学家长期以来持有的见解是:即使没有执行他们的确定所必需的操作,这样构造的概念仍然是"物自体"的固有属性。在量子理论中,我们面对着对这种形而上学观点的一种基本限制。
>
> (Weyl, 1928/1931: p.76)

最后,我列入了来自韦尔的一段引文,这个引文深受这类参考文献的相对频率问题的影响,不过,恰好是在量子力学中的同质性(homogeneity)问题的语境中作出的。

> 一般情况下,完成一个实验的条件甚至将不保证,构成被观察系统的所有个体都处于相同的"态",正如在量子理论中通过系统空间的一条射线所表示的那样。例如,情况正是如此:当我们只当心所有的原子都处于量子态(n, l)时,不能凭借斯特恩—盖拉赫效应相对于 m 把它们分离开来。为了应用量子力学,有必要设立这样一个标准:这个标准将使我们确定,为了保证这样的一个"纯态",已知条件

> 是否是充分的。我们说，如果(1) 在 $\mathfrak{I}$ 条件下有一个明确的、可重复的值的每一个量，在 $\mathfrak{I}'$ 条件下都有同样明确的值，如果(2) 存在着一个量，在 $\mathfrak{I}'$ 条件下是严格确定的，但在 $\mathfrak{I}$ 条件下并非如此，那么，条件 $\mathfrak{I}'$ 比条件 $\mathfrak{I}$ 产生了更大的同质性效果。这种渴望得到的标准显然是这样的：**如果不可能进一步增加同质性，条件 $\mathfrak{I}$ 就确保了一个纯态。**(在经典物理学中，只有当与系统相联系的所有的量都具有明确的值时，才能获得最大的同质性。)
>
> (Weyl, 1928/1931: pp. 77 - 78)

260

冯·诺伊曼。即使在今天，在许多方面对量子力学中的概率进行的最详细的讨论，也是在冯·诺伊曼(von Neumann)的经典论著(1932/1955)中找到的。概率特征的大多数分析集中于两个问题。第一个问题是，对比因果性过程和非因果性过程。按照依赖于时间的薛定谔方程，一个态随时间的演化在特征上是确定的和因果性的。另一方面，对处于离散光谱态的一个系统进行测量，可能经历变化到几个可能态的任何一个态。由于它的概率本性，量子力学中的这种测量过程是非因果性的。很显然，冯·诺伊曼在作出这种区分时，利用了经典物理学的严格决定论的因果性的观点。他提出的第二个问题与第一个问题密切相关，即将把概率的测量过程转化为一个经典的决定论的过程的决定论的隐变量的可能性。他给出一个证明：这样的隐变量是不存在的，现在认为这个证明是相当不能令人满意的。关于量子力学的隐变量的更新近的结果的概述在第 7 章给出，因此，这里不再追述。

在统计力学中比在量子力学中更引人注意的第三个问题是，量子过程的各态历经特征的问题，冯·诺伊曼对这个问题是有所贡献的，但是，他只在关于量子力学的书中提到它。也许，

因为他并不想在任何细节上把另一个概念层增加到当前的各态历经的思想中。在第四章关于不变性的最后一部分，给出了对各态历经过程的一个基本介绍，但没有触及对量子力学的应用。

在关于测量过程的最后一章，冯·诺伊曼写下了关于物理学中观察的主观性的最著名的一段，这在关于量子力学基础的大量文献中随处可见。我引用了它的重要部分。

> 首先，在本质上完全正确的是，测量或与主观感知相关的过程是相对于物理环境的一个新的实体，并不能还原为后者。确实，主观感知把我们带向了个人的内在的智力生活，根据其本性，这种智力生活是外在观察的（因为根据任何可想到的观察或实验，这一定是理所当然的）。（参见上面的讨论。）然而，它是科学观的一种基本的必要条件——所谓心理—物理平行主义的原理——为了描述主观感知的外在的物理过程，它一定是可能的，好像它实际上就在物理世界当中——也就是说，在客观环境中，即在日常空间中，把等价的物理过程赋予它的部分。（当然，在这个相关的程序中，产生出的常见的必然性是，把若干这些过程局域化在位于我们自己的身体占有的空间的那些点上。但是，这并没有改变这种事实：它们属于"关于我们的世界"，即上面所指的客观环境。）在一个简单的事例中，可以把这些概念应用如下：我们希望测量一个温度。如果我们需要，直到我们有了水银温度计的环境的温度时，我们才能追求用数字表示这个过程，然后说：这个温度是用温度计测出的。但是，我们能完成进一步的计算，而且，从能够用动能和分子的术语说明的水银的特性来看，我们能够计算它的加热、膨胀和水银柱的有效长度，然后说：这个长度是由观察者

261

看到的。还要继续进行下去,考虑光源,我们会发现,光量子在不透明的水银柱上的反射、其余的光量子到达观察者眼睛的路径、它们在眼睛透镜上的折射,以及视网膜上形成的图像,然后我们会说:这个图像是通过观察者的视网膜记录下的。如果我们的生理学知识比今天更精确,我们还能继续进行下去,在神经束中和在大脑中,追踪产生视网膜上的图像效果的化学反应,最后说:观察者感知到他的脑细胞的这些化学变化。但无论如何,不管我们计算进行多远——水银管、温度计的刻度、视网膜或进入大脑,在某个时候我们必须说:这是被观察者所感知到的。也就是说,我们必须总是把世界划分为两个部分,一部分是被观察的系统,另一部分是观察者。在前者的情况下,我们能够任意精确地追踪所有的物理过程(至少在原则上)。在后者的情况下,这是无意义的。两部分之间的边界在很大程度上是任意的。特别是在上面例子的四种不同的可能性中,我们看到,观察者在这种意义上不需要被等同于实际观察者的身体:在上面例子的一种情况下,我们连温度计都包括在内,而在另一种情况下,甚至没有包括眼睛和视神经束。这个边界能够被任意地深入推进到实际观察者的身体的内部,这一点是心理—物理平行主义原理的内容——但是,如果这种方法不是无意义地进行下去,即如果有可能与实验进行比较,那么就不会改变这样的事实:在每一种描述方法中,必须在某个地方划出这个边界。的确,经验只作出这类陈述:一位观察者进行了某种主观的观察;而且,决不像这样:一个物理量有一个特定的值。

(von Neumann, 1932/1955: pp. 418 - 420)

显而易见的是,尽管没有讨论关于主观概率的较广泛的文献资

料，但冯·诺伊曼对主观感知的强调，与主观概率的基本思想产生了共鸣。第二，他对心理物理平行主义的不成熟的评论提出了比描述概率的典型贝叶斯主义观点更基本的主观主义。

物理学中的实用主义。我们能轻易地推广物理学中关于对待概率态度的概述。但是，在运用概率的形式方法的意义上，如果不太关注概率的表征或解释的基本问题的话，实用主义就完全统治了概率在物理学中的运用。有一种观点是包括冯·诺伊曼在内的每一个人都认可的。这种观点是，在重复相同设计的实验时，通过被观察现象的相对频率，确证或检验量子力学的思想。上面从福克著作的第75页引用的陈述，几乎被每一位物理学家所认可。但是，关于实验的这种陈述，几乎是概率的每一种解释观都可接受的，也是从最极端的贝叶斯主义观点到最极端的倾向性观点可接受的。此外，随机性的描述方法、厌恶的概率的相对频率解释，在物理学文献资料中，即使提到，也不多见。更一般地说，在新近的一篇文章中(Suppes 1998)，我试图为现在与过去的大多数物理学的务实热情提供更加广泛的案例。①

262

统计实践。前面几段已经强调了关于物理学中对待概率本性的态度的实用性特征。足以令人惊讶的是，鉴于贝叶斯主义者与非贝叶斯主义者在统计学范围内的许多概念争论，关于统计实践中的实用主义的一种更加明确的态度，能够在关注错综

---

① 这样一种实用性态度统治了物理学，而且，应该统治经济学，卡特赖特(Cartwright，1999)已经有效地论证了这一点。另一方面，法恩(T. L. Fine，1973)的杰出著作，从不同的视角涵盖了本章的许多部分，这本书关注对在更深层次上令人满意的概率表征的持续需要。在物理学家的实用主义表达的观点中的矛盾冲突，以及法恩为更大的和更特殊深刻的解释与实践的辩解，是科学方法的哲学基础的核心。由于这里无法详细地讲清楚的理由，我预计，物理学家的最低论者的解释纲领将继续统治概率概念在科学理论中的运用，这是为什么我在已经过分长的一章中还包含了最后这一部分的原因所在。

复杂的细节问题的统计学家中间找到。我将不试图概述这方面的文献资料,而是引证一个精彩的事例,这个例子完全不是来自物理学和量子力学,而是来自比我能想到的几乎任何一个量子力学事例都更加精致的统计的观点。

我想到的这个例子是,《联邦党人文集》的统计分析是确定,某些有争议的部分的作者,究竟是汉密尔顿(Hamilton),还是麦迪逊(Madison)。这本"文集"是由美国革命时期的一系列成文的最有名的政治文件组成的。仍然有许多人阅读这些文件,而且,有争议的原作者一直是一个富有历史兴趣的问题。莫斯特勒(Mosteller)和华莱士(Wallace)通过极其详细的统计研究,永远明确地解决了这个问题,即使以后还会在统计与计算方法上有所改进。下面是他们在第一版的序言中所表达的怀疑主义的实用性态度。

> 有些人可能觉得,我们应该对不同推理方法的比较价值——特别是贝叶斯式的推理方法与经典的推理方式——作出更明确的声明。显然,我们更喜欢倾向于贝叶斯式的进路,不过,至少在数据分析的领域内,分析技巧还很不成熟。随着这些技巧的演进,判断它们的优缺点会更容易。
>
> 即使个别统计学家在一般情况下可能要求遵循贝叶斯学派或经典学派,但是,没有一个人有足够的准则确定,在特定的时刻代表了什么学派。当我们以为我们是最好的贝叶斯主义者时,经典学派的人却告诉我们,我们完全是经典的;而当我们以为我们自己给出了经典的处理时,贝叶斯主义者却告诉我们,这种思想并不在经典的词汇当中。因此,我们除了自己之外,不能自称讲给任何人。
>
> (Mosteller and Wallace, 1964/1984: p. ix)

自莫斯特勒和华莱士于1964年第一次发表这篇文章以来，贝叶斯的概念和方法得到了广泛的发展，但事实上，在科学文献中，对贝叶斯方法的运用，仍然相对少见，尽管近来有像鲁阿内等人(Rouanet, 1998)撰写的如此出色的文本。但这并非因为科学家致力于客观的概率观，而不是主观的概率观——用莫斯特勒和华莱士的术语来说，统计学家的“经典学派”支持客观的概率观。不，这是因为基本上科学家在大多数学科中仍然不关心概率的基础问题，而且，在实用性的意义上应用统计概念，毫无顾虑地只计算某人教他们运用的统计检验。但另一方面，当需要对卷入很强实践的某些过程或现象作出统计分析时，不关心有时就消失了。 263

纯科学本性的其他力量也能改变态度，这最有可能发生在整个20世纪。统计学自身现在正处于革命当中，这种革命比统治20世纪下半叶的客观概率观和主观概率观之间的矛盾冲突有更深远的启迪。分析非常大的数据库所要求的新的统计方法，拉开了这场革命的序幕，现在，这场革命几乎在科学的每一个领域内都很流行。已经提出的这种方法在本性上是计算的，而且，无疑，将会继续如此。关于这个新的语境和焦点有许多讨论，讨论更多与物理学相关，较少与五十年前的统计学相联系的数据的科学来源相关。物理学家还有某些最好的统计实践例示的如此好的实用性态度，在相当长的时间内，将是对待概率本性的占有优势的态度，相信这种观点是合理的。从基本的观点来看，我不由得要看一下，第6节为了分析客观倾向解释所用的定性的公理化方法，和第7节为了分析概率的主观解释所用的那些方法，在基本意义上，有多少是相同的，而且，回想起来，我明白，这些方法可能通过更统一的形式结果得到改进。

这导致了我乐意承认的那种实用主义。统一的形式方法伴

随着对所用概念的非常不同的解释或表征。概率在应用中太丰富和太多样,以至于不能只限于一种过于自信的表征。运用的实用语境有时将确定所选择的表征,但经常一种更深层的实用态度将占有统治地位,而且将对表征作不出明确的选择,前面提到的莫斯特勒和华莱士(1964/1984)的著作中,以及最近在关于数据挖掘、自适应统计、推进(boosting)算法、神经网络和大规模数据分析的各种其他进路方面纷至沓来的文章中,在统计意义上,充分地例证了这种观点。[①]

---

① 关于这些最近发展的一个好的不太技术性的概述,可参见 Hastie, Tibshirani and Friedman (2001)。

6. 空间和时间的表征

早在亚里士多德作出广博而复杂的关于世界是永恒的、无始无终，并且只有一个世界（《天论》第一卷）的著名论证之前，对空间和时间的各种表征就是古希腊哲学中的一个重要论题。这种强调持续了许多世纪。也许，根据其时代的详细的物理学语境，把空间和时间的表征看成是他的哲学核心的最后一位主要哲学家，是撰写了《纯粹理性批判》和《自然科学的形而上学基础》的康德。无论如何，对空间和时间进行详细的概念分析，不再是大多数哲学家的中心论题，而是具有专业知识并对物理学基础感兴趣的那些人的一种盛行的活动。本章的前三节关注后面的这种类型。第一节陈述了一些必要的几何学的基本知识。第二节假设，在经典物理学中运用了标准的解析几何表征，并集中研究了这种表征的不变特性，但为了证明这种不变性，只需要经典力学的运动学特性。第三节对狭义相对论的时空进行了类似的分析。这种分析反映了在经典物理学和相对论物理学中不变性问题的重要性；在这两节的最后作出了某些相当广泛的历史评论。用到的非常简化的假设，不是根据更主要的物理学概念建立起来的，这个假设是，经典时空结构和受限制的相对论的时空结构都是第一节定义的仿射结构的特殊情况。 265

对这种仿射统一性有几种辩护。首先,经典时空和狭义相对论时空的许多特性在一般的仿射变换群下都是不变的——例如,惯性轨道是线性的并且一直保持平行。更重要的是,经典物理学的伽利略变换群和狭义相对论的洛伦兹变换群都是仿射群的子群。这些共同特征通常在空间和时间的哲学讨论中,没有给予足够的重视。

在第四节,关注的焦点从物理学转向心理学,特别是转向这样的问题:我们如何确定视觉空间是否是欧几里得的。因为哲学家较少考虑空间知觉现象,所以,这部分的论述会更加详细。如期所料,这些结果在特征上是不太明确的。继续沿着这一思路,第五节考察了关于视觉空间本性的各种哲学主张和经验断言。这一节较详细地讨论的福莱(Foley)和瓦格纳(Wagner)的心理学实验在第 6 节提供了某些受限制的全等公理的框架。这些公理扩展了定义 1 的纯仿射公理,但避免了由福莱和瓦格纳的实验产生的曲率不变的任何空间(不只是欧几里得空间)的悖论。第七节随着对视觉现象的任何一个空间理论的三个主要问题的反思,结束了对视觉空间的考察。

266

在最后一节,即第八节,我简要地转向数学哲学中的一个相关主题:几何学中的有穷论。从概念上来看,所考虑的是几何作图的有穷性和用这种方法能够解决的那类问题,几何作图的整个悠久传统可追溯到最早期的古希腊几何。

正如能够看到的那样,这一章的不同章节从物理学转到心理学,最后再到几何学本身。这似乎有点像混合物,确实如此。第二节和第三节关于物理学中的不变性原本可以被放在下一章(第七章)的开头,即关于力学中的表征。但是,由于三种理由,我还是把它们放在这里。第一,在这两节没有正式用到动力学假设,因此,力学思想的使用相当薄弱。第二,洛伦兹变换的推

导更多地与电磁学理论相关，特别是在历史上不太与力学相关。第三，我在这一章想强调关于空间和时间表征的许多不同思考方式。关于空间和时间现象的不同学科提出了需要用不同的进路来解决的问题。

6.1 几何学的基本知识

本章提出的对待时空的观点是，标准的几何概念和对象是被假定的，而且没有得到详细的阐述。事实上，所需要的是一个仿射空间的概念和一个欧几里得空间的概念，以及用来作出这些空间的像介中性之类的几何关系。

为了保持对经典时空或明可夫斯基时空所需要的特殊公理的首先关注，我只会简要地评论仿射空间和欧几里得空间的一些标准的和熟悉的公理。我将阐述表征和不变性定理，以备后用。

在一种标准的版本中，依据非空集 A 和介中性的三元关系 B，把**有序的仿射空间**公理化。[我也用标准的简化符号 $a|b|c$ 表示 $B(a, b, c)$。]在这些空间的坐标表征中，点是实数的有序的 n 元组，形成了标准的 n 维笛卡儿空间 R^n。设 $\mathbf{x}$、$\mathbf{y}$ 和 $\mathbf{z}$ 是任意三个这样的笛卡儿点。于是，在预期的解释中，介中性 $B_R(\mathbf{x}, \mathbf{y}, \mathbf{z})$的笛卡儿关系成立，当且仅当，存在一个 $\alpha \in [0, 1]$，使得

$$\mathbf{y} = \alpha \mathbf{x} + (1-\alpha)\mathbf{z}$$

结构(R^n, B_R)是 n 维的仿射笛卡儿空间。

我们需要下列定义。对于 $a \neq b$，一条**直线** ab 是所有的点 c 的集合，使得如果点 a、b 和 c 不共线，即不在一条直线上，那么 $a|b|c$、$b|c|a$ 或 $c|a|b$ 和 abc 是一个**三角形**。(注意，这里的一

个三角形只是三个不共线的点的一个有序三元组。由于符号的简化,用同样的符号表示由点的这样一个三元组所定义的点集构成的平面。)

267

设 a, b, c 是不共线的;于是,平面 abc 定义如下:

$$abc = \{d \mid (\exists e)(\exists f)(e \neq f,\text{并且},e,\ f \in ab \cup ac \cup bc\ \text{并且},d、e、f\ \text{是共线的})\}$$

(通过要求 $e|d|f$,而不只是共线,我们定义了三角形区域 abc。)对于三维空间,我们用了由不共面的四个点组成的**四面体** $abcd$。下列常用的术语 a、b、c 和 d 是这个四面体的四个**顶点**,线段 ab、bc、ac、ad、bd 和 cd 是**边**,四个三角形区域 abc、bcd、cda 和 dab 是**面**。三维空间 $abcd$ 就是与在四面体 $abcd$ 的一面或两面的点对(pairs of points)共线的所有点的集合。

我们以同样的方式继续到四维空间,即经典空间和狭义相对论的空间所需要的一个概念。五元组 $abcde$ 是一个**单纯形**,当且仅当,五个点都不在同一个三维空间里。这个单纯形的**胞腔**是由单纯形的四个点决定的五个四面体的区域。四维空间 $abcde$ 就是与这个单纯形的一个或两个胞腔上的点对共线的点集。

定义 1. 一个结构 $\mathfrak{A} = (A, B)$ 是一个仿射空间,当且仅当,A 中的 a、b、c、d、e、f、g、a'、b' 和 c' 都满足下列公理:

1. 如果 $a|b|a$,那么 $a = b$。

2. 如果 $a|b|c$,那么 $c|b|a$。

3. 如果 $a|b|c$ 和 $b|d|c$,那么 $a|b|d$。

4. 如果 $a|b|c$ 和 $b|c|d$ 且 $b \neq c$,那么 $a|b|d$。

5. (连通性)如果 $a|b|c$、$a|b|d$ 且 $a \neq b$,那么 $b|c|d$ 或 $b|d|c$。

6. (外延性) 在 A 中存在 f,使得 $f \neq b$ 和 $a|b|f$。

7. [帕施公理(Pasch's Axiom)]如果 abc 是一种三角关系,$b|c|d$ 和 $c|e|a$,那么在直线 de 上存在着一个点 f,使得 $a|f|b$。

8. (完备性公理)对于一条直线到两个非空集 Y 和 Z 的每一种划分,使得

(i) Y 的点 b 不在 Z 的任意两点 a 和 c 之间,而且,

(ii) Z 的点 b' 不在 Y 的任意两点 a' 和 c' 之间。

存在着 $Y \cup Z$ 的一个点 b,使得对于 Y 中的每个 a 和 Z 中的每个 c,b 都在 a 和 c 之间。

9. (维数)。此外,如果 $\mathfrak{A}$ 的所有点都在一个平面上,但不在一条直线上,那么 $\mathfrak{A}$ 是二维的;如果 $\mathfrak{A}$ 的所有点都在一个三维空间里,但不在一个平面上,那么 $\mathfrak{A}$ 是三维的;如果 $\mathfrak{A}$ 的所有点都在一个四维空间里,但不在一个三维空间里,那么 $\mathfrak{A}$ 是四维的。[①]

只根据介中性的这种阐述是相当复杂的,因此,我通过运用在定义 1 之前定义的概念,回避这种阐述。[②]

定理 1. 对于 $n = 2$、3、4,每个 n 维仿射空间(A, B)与 n 维仿射笛卡儿空间(R^n, B_R)都是同构的,即存在着把从 A 映射到 R^n 的一个一一对应的函数 f,使得对于 A 中的任意点 a、b 268

---

① 公理 9 是一种定义的形式,但它给出了确定维数所需要的公理形式。例如,如果某人想关注二维的仿射空间,维数公理是:$\mathfrak{A}$ 的所有点都在一个平面上,但不在一条直线上。

② 我注意到,仿射几何的一种完全不同但等价的进路是从一个射影空间开始的,比如说,以三维空间为例,然后,去掉贴上"无穷远平面"标签的一个平面。结果得到的空间是仿射的。在二维的情况下,去掉一条无穷远的直线,得到了这个仿射平面。

和 c,

$$B(a, b, c), \text{当且仅当}, B_R(f(a), f(b), f(c))$$

此外,任意两个 n 维仿射空间是同构的。

换言之,这个定理断言了一个已知维度的仿射空间的范畴特征。(后面在第 8 节引入的有穷仿射作图的基本公理不是绝对的。)

相伴随的关于表征的不变性或惟一性定理是同等重要的,尽管有些人忽略了这个问题。用标准的几何术语来说,不变性只惟一地等于仿射变换群。一种仿射变换是 n 维笛卡儿空间的任意一种线性变换 φ,使得对于每一个向量 $\boldsymbol{x}$,①

$$\varphi(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b},$$

这里 $\boldsymbol{A}$ 是一个非奇异的 $n \times n$ 矩阵,$\boldsymbol{b}$ 是任意一个 n 维向量,即实数的任何一个 n 元组。从直观上看,矩阵 $\boldsymbol{A}$ 表示一种旋转,同时,沿着 n 维笛卡儿坐标延长不同的量。向量 $\boldsymbol{b}$ 表示笛卡儿空间的原点的一种平移。关于这些变换的两个熟悉的事实是:它们保持共线,以及它们从平行线变换为平行线。仿射变换的集合,在函数复合条件下,也显然是一个群。

定理 2. 根据在实数域上的 n 维仿射笛卡儿空间,一个 n 维仿射空间的表征,在仿射变换群下,是不变的。特别是,如果 f 和 f' 是把已知的仿射空间同构地映射到 R^n 的两个函数,那么存在着一种仿射变换 φ,使得 $f' = \varphi \circ f$;同样,如果 f 是把已知的仿射空间如此映射到 R^n 的一个函数,φ 是任意仿射变换,那

---

① 本章的前三节改变了在第 5 章中约定的符号。黑体的大写字母,比如 $\boldsymbol{X}$ 或 $\boldsymbol{A}$,不再指随机变量,而是指矩阵,而且现在,黑体的小写字母,比如 $\boldsymbol{x}$ 和 $\boldsymbol{y}$,是指向量。

么，$\varphi \circ f$ 也是这样一个同构映射。

正如后面将明确说明的那样，伽利略变换和洛伦兹变换是两种特殊的仿射变换。

为了从仿射空间获得欧几里得空间，我们只需要增加全等或等距的四元关系：$ab \approx cd$，当且仅当，a 点和 b 点之间的距离等同于 c 点和 d 点之间的距离。用不同的术语来说，以 a 和 b 为端点的线段全等于以 c 和 d 为端点的线段。对笛卡儿空间 R^n 中的等距解释是熟悉的：对于 R^n 的任意向量 $\mathbf{x}$、$\mathbf{y}$、$\mathbf{u}$ 和 $\mathbf{v}$，关系 $\mathbf{xy} \approx_R \mathbf{uv}$ 成立，当且仅当，

$$\sum_{i=1}^{n}(x_i - y_i)^2 = \sum_{i=1}^{n}(u_i - v_i)^2$$

我省略了熟悉的欧几里得空间的全等公理，有必要把这些公理增加到定义 1 的仿射公理中。它们推广了仿射只对平行线成立的全等概念。潜在地假设这些全等公理，于是，我们有标准的表征： 269

定理 3. **每一个 n 维的欧几里得空间 $(A, B, \approx)$ 都同构于 n 维的欧几里得笛卡儿空间 $(R^n, B_R, \approx_R)$，即存在着把 A 映射到 R^n 的一个一一对应的函数 f，使得**

(i) 对于 A 中的任何 a、b 和 c 点，$B(a, b, c)$，当且仅当，$B_R(f(a), f(b), f(c))$

和

(ii) 对于 A 中的任何 a、b、c 和 d 点，$ab \approx cd$，当且仅当，

$$\sum_{i=1}^{n}(f_i(a) - f_i(b))^2 = \sum_{i=1}^{n}(f_i(c) - f_i(d))^2$$

这里 $f_i(a)$ 是笛卡儿向量 $f(a)$ 的第 i 个坐标。此外，任意两个 n 维的欧几里得空间是同构的。

另一方面,定理 3 伴随的不变定理,是同样重要的。正如上面定义的那样,一个仿射变换 $\varphi(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x} + \boldsymbol{b}$ 是一个相似变换,当且仅当,矩阵 $\mathbf{A}$ 是一个相似矩阵,即存在着一个正交矩阵 $\mathbf{A}'$ 和一个正实数 δ,使得 $\mathbf{A} = \delta\mathbf{A}'$。换言之,一个相似变换只允许在所有方向上的一个统一的单位延长,反而一个一般的仿射变换没有强加这种限制。

定理 4. **根据实数域上的 n 维欧几里得笛卡儿空间,一个 n 维的欧几里得空间的表征,在相似变换群下,是不变的。**

与前面的陈述相反,一种任意选择的伽利略变换或洛伦兹变换,在它的空间部分,不是一个一般的相似变换,因为空间距离的测量单位保持不变,这是显然的,但却是值得注意的。这一点会在下面两节变得更清楚。

6.2 经典时空

如果不是给出一种纯定性的几何表述,就像在苏佩斯等人的著作(1989,第 13 章)中可以看到的关于这些定性表述的历史那样,引入参照系,即笛卡儿正交坐标系中的一维时间坐标和三维空间坐标,在物理学上也是很自然的。

从形式上看,一个时空测量框架 f 是把每个时空点 a 都映射到笛卡儿空间 $(R^3, B_R \approx_R)$ 里的一个空间向量$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ 的一个函数对(s, r),这里,在空间上,$s_i(a) = x_i$,在时间上,$r(a) = t$。相对于一个参照系表示测量的这些空间函数和时间函数,对于经典物理学和狭义相对论物理学,是相同的。当我们从经典的情形过渡到相对论的情形时,实验室或测量的观察方法没有发生改变,但基本的不变量却发生了改变。在下面给出的公理中,我们参照了惯性系,但只假定这些惯性系有仿射性和

适当的不变性。[①] 任意两个惯性系通过不变的相对速度联系起来，是一种结论，而不是支持不变性公理的一个假设。 270

定义 1. 一个结构 $\mathfrak{A}=(A, B, \mathcal{F})$ 是一个惯性系的四维经典时空结构，当且仅当，满足下列公理：

公理 1. 结构 (A, B) 是一个四维仿射空间。

公理 2. $\mathcal{F}$ 是把 (A, B) 同构地映射到仿射笛卡儿空间 (R^4, B_R) 的（惯性）系 $f=(s, r)$ 的一个非空集。

公理 3.（时间不变性）对于 A 中的任意两点 a 和 b 以及 $\mathcal{F}$ 中的任意两个惯性系 $f=(s, r)$ 和 $f'=(s', r')$，时间间隔是不变的，即

$$|r(a)-r(b)|=|r'(a)-r'(b)| \tag{I}$$

公理 4.（空间不变性）对于 A 中的任意两点 a 和 b 以及 $\mathcal{F}$ 中的任意两个惯性系 (s, r) 和 (s', r')，空间距离是不变的，即

$$\sum_{i=1}^{3}(s_i(a)-s_i(b))^2=\sum_{i=1}^{3}(s'_i(a)-s'_i(b))^2 \tag{II}$$

公理 5.（不变的惯性系的闭合性）如果 $f=(s, r)$ 在 $\mathcal{F}$ 中，如果 φ 是 (R^4, B_R) 的一个仿射变换，使得 f 和 $\varphi\circ f$ 满足(I)和(II)，那么，$\varphi\circ f$ 在 $\mathcal{F}$ 中。

这些公理中的每一个公理的直观意义应该都是清晰的，但是，作出某些评论可能是有用的。在这里给出的分析中，经典时

---

① 从物理学的观点来看，令人向往的是，像经典用语所说的那样，假定存在着一个“相对于固定恒星是静止的”固定参照系。但是，引入这种假定显然会打乱定义1中给出的公理化描述的纯形式特征。然而，在物理学上重要的直觉是，一个惯性系是以一种不变的速度相对于某个参照点运动的系统，这个参照点具有巨大的或永恒的特征，对于牛顿来说，是绝对空间，对于一些人来说，是固定的三维坐标，即遥远的恒星，对于另外一些人来说，是太阳系的质心，或者，最终用更宏大的术语来说，是宇宙的质心，在以一种接近不变的速度作相对运动的意义上，后面三种说法显然在经验上是非常相似的。

空和相对论时空共有的公理 1 表达了空间的仿射特征。它们共有的最重要的特征之一显然是,这两种空间都是平坦的或线性的。一个仿射四维时空结构的假设,在经典物理学的时间和空间的基本分析中,不是常用的。通常是对时间和空间进行单独分析,但是,这里所用的统一性自然只解释了,经典时空与限于相对论的时空如何取决于共同的假设,并且它们如何不同。从一种定性的几何观点来看,如果我们提出物理空间和时间的一条完全定性的进路,那么,只根据介中性公理化的仿射框架,就必须补充额外定性的基本概念。在同时发生的两个点事件(point events)之间的空间距离的一个全等概念,对经典时空来说,自然是附加的。然而,物理学家几乎总是回避这条定性的几何进路,即使在基本的讨论中,也是如此。这里所用的惯性系更加接近于物理学的思维方式。假定经典时空和相对论时空具有一个共同的四维仿射空间,对这种假定的另外一种辩护是,两者共有的一种不变性是,质点的惯性轨道都是线性的,即质点受到的合力为零的时空轨迹。

271

我们不要求为经典时空结构提出一个独立的表征定理,因为表征函数恰好是在定义 1 的公理 2 中假设的惯性系。所需要的是伽利略的不变性定理。

经典时空结构的不变性定理是定理 6.1.2 关于仿射空间表征的不变性的一种特殊化。回忆一下,惟一性相当于是一个任意的仿射变换。我们现在把一个附加条件强加于这种变换的矩阵 $\mathbf{A}$。我们说,一个四维矩阵 $\mathbf{A}$ 是一个相对于 $\mathcal{F}$ 中的惯性系 f 的伽利略矩阵,当且仅当,存在一个三维向量 $\boldsymbol{u}$、一个 3 阶正交矩阵 $\boldsymbol{E}$ 和 $\delta=\pm 1$, 使得

$$\mathbf{A}=\begin{pmatrix}\boldsymbol{E} & 0\\ 0 & \delta\end{pmatrix}\begin{pmatrix}\boldsymbol{I} & -\boldsymbol{u}^{*}\\ -\boldsymbol{u} & 1\end{pmatrix}$$

这里 $\boldsymbol{I}$ 是 3 阶单位矩阵。

常数的解释如下。数字 δ 表示在时间方向上的一种可能变换——如果 $\delta=1$，那么时间方向不变，如果 $\delta=-1$，那么时间方向相反；$\boldsymbol{u}$ 是新的惯性参照系相对于旧参照系的相对速度；矩阵 $\boldsymbol{E}$ 表示空间坐标的一种旋转。仿射变换的向量 $\boldsymbol{b}$ 也有一种简单的解释。前三个坐标表示 f 的空间坐标原点的平移，第四个坐标表示 f 的时间测量的一维坐标系的原点的平移。我们因此假定，在 $\mathcal{F}$ 中的所有不同的惯性系，都用了同样的空间和时间测量尺度，正如在公理 3 和 4 中所蕴含的那样，但是，当从一个惯性系运动到另一个惯性系时，允许空间或时间位置的平移。

我们说，一种仿射变换是**伽利略的**，当且仅当，它的四维矩阵是伽利略的。运用这些概念，我们就可以陈述经典空间和时间的下列不变性定理。

定理 1. **根据实数域上的四维仿射笛卡儿空间，一个经典时空结构的表征，在伽利略变换群下，是不变的。**

证明： 设 f 和 f' 是 $\mathcal{F}$ 中的两个惯性系。我们分析仿射变换 φ，使得对于 f 的每一个四维笛卡儿向量 $\boldsymbol{x}$，$\varphi(\boldsymbol{x})=\boldsymbol{x}'$。由于定理 6.1.2，所以，存在一个 4 阶非奇异矩阵 $\boldsymbol{A}$ 和一个四维向量 $\boldsymbol{b}$，使得对于 f 的每一个四维笛卡儿向量 $\boldsymbol{x}$，

$$\varphi(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}$$

这个证明变成了表明 $\boldsymbol{A}$ 是一个伽利略矩阵。既然 $\boldsymbol{A}$ 是非奇异的，因此，我们可以把 $\boldsymbol{A}$ 分解为

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{D} & \boldsymbol{e}^{*}\\ \boldsymbol{h} & g\end{pmatrix}$$

这里 $\boldsymbol{D}$ 是一个 3×3 的非奇异矩阵，代表了 f 的任何一个向量 $\boldsymbol{x}$

的三维空间坐标 x_1、x_2、x_3，$\boldsymbol{h}$ 是一个行向量，$\boldsymbol{e}^*$ 是一个列向量，272 $g \neq 0$ 是一个实数。

首先，假设 $\boldsymbol{D}$ 不是一个正交矩阵。于是，很容易表明，存在着向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，使得与公理 4 相反，不能通过 φ 保持 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的空间距离，即

$$\sum (x_i - y_i)^2 \neq \sum (\varphi(x_i) - \varphi(y_i))^2$$

根据类似的论证，我们能够表明，由于定理 3，所以，$g = \delta = \pm 1$。

关于经典力学的某些一般定理可以在 7.1 节中找到，但是，这些定理一定超出了本节的运动学框架的范围。

历史上的评论。对于不熟悉经典力学基础的读者来说，定义 1 的形式的不变性公理，乃至提供一般框架的仿射空间上的第一公理，直观上是不明显的，这些思想起源于哪里，是从哪里开始它们被从其力学语境中抽象出来而具有几何的性质。惟一留下的但当然是基本的经典物理学的剩余是，一个时空点的坐标定位的前三个坐标是物理空间意义上的空间坐标，一个时空点的第四个坐标定位是物理时间意义上的时间坐标。

这个不变性的发现是相对较晚的。它无疑与亚里士多德的物理学不相符，这种物理学认为，宇宙的中心就是地球的中心。更重要的是，从观察推断出，天体(即，行星和恒星)的自然统一的运动是圆周运动。

下面是托勒密对这种观点的有力概括：

> 一般的初步讨论涵盖了下列论题：天体在形状上是球形的，像一个球体一样运动；当把地球看成一个整体时，地球在形状上显然也是球形的；在位置上，它位于天体的中间，很像天体的中心一样；在大小和距离上，它具有一点与

固定恒星球面的比率;它不会从一个地方运动到另一个地方……似乎合理的是假设,古人从下列类型的观察获得了他们关于这些论题的最初的概念。他们看到,太阳、月亮和其他星体沿着圆周东升西落,这些圆周总是彼此平行的,好像它们从地球下面开始升起,逐渐升高,然后,以相似的方式保持圆周运动,直到逐渐地下落到低于地球为止,可以这么说,它们完全消失了,然后,在看不见一段时间之后,又重新升起和落下;并且[他们看到]这些[运动]周期,还有升起与落下的地方,基本上是固定的和相同的。

致使他们形成球体概念的主要原因是,始终能看到的恒星的旋转,所观察到的旋转是圆周运动,并且总是围绕一个中心运动,[所有的]星体都是如此。由于不可避免的原因,[对他们来说]那个点就成为天球的极:靠近它的那些恒星围绕较小的圆周旋转,远离它的那些恒星围绕所描述的更大的圆周旋转,圆周的大小与它们的距离成比例,直到恒星遥远得超出了人们可视范围为止。而且在这些情况下,他们也看到,那些近的总能看见的恒星在短时间内仍然会看不见,而那些遥远的恒星在长时间内总是一直看不见,还是[与他们的距离]成比例。结果是,他们从一开始就完全根据这些考虑得到了前面提到的概念;但从此之后,在他们后面的研究中,他们发现,其他一切都与此相符,因为所有的现象都肯定与所提出的其他替代概念相矛盾。 273

因为如果一个人假设,像有些人认为的那样,星体是在一条直线上开始运动,直到无穷远,那么他能想象得出是什么原因导致每个星体似乎每天都从相同的起点开始他们的运行呢? 如果这些星体运动到无穷远,那么它们如何能够返回呢? 或者,如果它们确实能返回来,这一点为何不能是

明显的呢？[根据这样一种假设]，它们一定在尺度上逐渐地变小直到消失为止，但却相反，在它们消失的那一瞬间，看上去是更大了，这时，它们好像是被地平面阻隔和挡住了……

总而言之，如果一个人假定，天体的运行是除了球形之外的任何其他形式，那么，他设想，无论把地球本身定位在哪里和无论如何定位，都必然会断定，所测出的星体离地球表面的距离一定是可变的。因此，对于相同的观察者来说，在每次旋转的过程中，星体的大小和彼此间的距离一定似乎是可变的，因为有时它们一定在更远的距离，有时在另一个距离较近的地方。然而，我们看到，根本没有发生这样的变化。

(Ptolemy, *Almagest*: pp. 38 – 39)

上面所强调的观点是，关于匀速圆周运动的这个结论，不是来自一个先天的论证，①而是根据几百年来记录的天文观测得出的一个推论。已知这种伟大的成功和相当的精确性，难怪花费了数百年的时间，才推翻了亚里士多德的这些结论。

此外，根据地球上的现象，而不是根据天文观测，提出新的思想，是毫不奇怪的。在中世纪的力学中，除了动力概念的长期发展之外，贝内代蒂(Benedetti, 1585/1959)也许是坚持认为自

① 著名的物理学家大卫·博姆(David Bohm)在下列一段引文中(1965, p. 5)全力支持了这种错误的先天观点，这段引文取自他关于狭义相对论的教科书中：

亚里士多德学说的一个组成部分是，比地球上的物质更完美的天体(比如行星)应该沿着一条表达它们的完美本性的轨道运动。既然圆被认为是最完美的几何图形，所以，可以断定，一颗行星一定围绕地球作圆周运动。当观察未能揭示出这种完美的圆形时，通过引入"本轮"或"圆中圆"，来调节这种不一致。这样，托勒密的理论得到了发展，这种理论以一种非常复杂的方式引入许多本轮，能够使自己适应于任何一条无论什么样的轨道。

由运动的物体只沿着一条直线运动的第一人。他关于旋转的抛物体是这么说的：

> 的确，当物体旋转时，圆周运动使得物体由于自然的倾向性(naturalis inclinatio)和在初始推动力的作用下沿着一条直线轨道(recta iter peragere)运动。
>
> (Benedetti, in Clagett, *The Science of Mechanics in the Middle Ages*, p. 664)

只不过不久之后，伽利略自己把惯性运动想象为是直线运动的进步，是缓慢的和复杂的。①

这里无法详细地概述17世纪的快速发展，但几乎等价于牛顿第一定律的一种早期的明确表述可以在笛卡儿的《哲学原理》第二卷(1644/1983, pp. 59－60)的命题37和39中找到： 274

> 37. 第一自然定律：每一种东西，就其能力而言，总是保持相同的状态；结果，它一旦开始运动，总是一直运动。
>
> 39. 第二自然定律：所有的运动在本质上都是沿着直线进行的；因而，作圆周运动的物体总是倾向于远离它们正在描绘的圆的中心。

为提出经典的不变性概念扫清道路，需要的正是惯性运动的这个明确的直线概念，还有对固定不动的中心点的拒绝。②

无疑，笛卡儿强调运动的相对性的一个原因是，为了在《哲

---

① 例如，参见在中世纪力学背景下对克拉格特的讨论(Clagett 1959, pp. 666－671)。

② 笛卡儿在第二卷的命题13的讨论中，也在其他地方，都拒绝接受不动的中心。

学原理》的第三卷中，他能够运用哥白尼的体系，但仍然不要求地球运动。下面是第三卷的命题 19(1644/1983: p. 91):

> 19. 我比哥白尼(Copernicus)更谨慎而且比第谷(Tycho)更真实地否认地球的运动。①

在科学和天主教神学之间的矛盾冲突的难点上，笛卡儿是一位老练的和事佬。

尽管牛顿在他的《自然哲学的数学原理》(1687/1946, p. 9)的开始几页提出了绝对静止的概念，但是，他以一种谨慎小心的方式断言，"绝对空间无法根据在我们领域内的物体的位置来确定"。他继续阐述了相对运动概念和真实运动概念的用法和需要。

但是，在经典物理学中，完全承认和运用不变性，还需要很长时间。我认为，在马赫(1883/1942)的重要但有缺陷的基础研究《力学及其发展的批判历史概论》一书的九个版本中，没有一个版本系统地考虑过不变性的问题。同样，惠特克(E. T. Whittaker, 1904/1937)的分析动力学的名著，一本完全不同特征的著作，也是如此。

事实上，这很像是时间的一种回溯归纳。我们不得不引入首先强调涌现出来的不变性的爱因斯坦的狭义相对论，作为一种自然方式说明与经典物理学的不同之处。只是在这个后来的语境中，才能完全地和系统地阐述仿射假设和定义 1 的不变性，包括对伽利略变换的明显认同。②

---

① "第谷"是指第谷·布拉厄(Tycho Brahe, 1546 - 1601)，他以对天文学的出色的系统观测而著名，并且对于开普勒(Kepler)的理论工作具有很大的重要性。

② 根据亚伯拉罕·派斯(Abraham Pais, 1982, p. 140)的观点，弗兰克(Frank, 1909, p. 382)首先提出了"伽利略的"这个术语。

6.3 狭义相对论的公理

本节的目的是描述狭义相对论的不变特性。与上一节的情形一样,把空间的几何点解释为四维的点事件(point-events)。根据洛伦兹变换群,给出所期望的不变结果。 275

像在经典情形中一样,线性是借助于介中性的三元关系引入的。这里所用的进路会使线段的全等问题,即等距离问题,保持在最显著的位置。当然,对距离的空间测量和时间测量的不变性分别进行简单的经典表述是不可能的。无论如何,正如许多年前罗布(1936)指出的那样,在狭义相对论中,对于惯性线的线段,即匀速运动的质点的可能路径来说,存在着一个有意义的全等概念。这个全等是惯性线段的**固有时间**的全等。对于分隔的线段或空间线,即对于既不能确定一条惯性线也不能确定一条光线的两点所确定的那些线段来说,同样有一个不同的全等概念。相反,对于光线(即,表示光路径的直线)的线段来说,根本没有自然的全等概念。

通过引入狭义相对论的通常的坐标表示法,最容易说明对这些思想的直觉支持。正像在经典物理学中那样,相对于一个惯性时空参照系 f 来测量一个点事件的四个实坐标(x_1, x_2, x_3, t)。像上一节那样,坐标 x_1、x_2 和 x_3 是三个正交的空间坐标,t 是 f 的时间坐标。为了符号上的方便,常用 x_4 代替 t。两点 $x=(x_1, x_2, x_3, x_4)$ 和 $y=(y_1, y_2, y_3, y_4)$ 之间的相对论性的距离或**固有时间**,相对于参照系 $f=(s, r)$,通过不定的二次度量被规定为

$$I_f(xy)=\sqrt{c^2(x_4-y_4)^2-\sum_{i=1}^{3}(x_i-y_i)^2} \qquad (1)$$

这里 c 是光速的数值。更加简单的是运用相对论性距离的平方。[①] 容易看出,如果 x 和 y 在一条惯性线上,那么 $I_f^2(xy) > 0$; 如果 x 和 y 在分隔线上,那么 $I^2(xy) < 0$; 如果 x 和 y 在一条光线上,那么 $I^2(xy) = 0$。在这些术语中,两条惯性线段 xy 和 uv,即惯性线的线段是全等的,当且仅当,

$$I_f^2(xy) = I_f^2(uv) \tag{2}$$

而且,类似条件对分隔线的两条线段成立。另一方面,这个等式对光线根本不起作用,因为任何一条光线段的相对论性的长度都是 0。

已知一个仿射时空结构,真空中光速 c 的不变性和不同惯性系的惯性线段的固有时间的不变性,是证明惯性系必定通过洛伦兹变换相关联的一个充分条件。限于惯性线段,在物理意义上,是自然的,因为这样的线段比分隔线的线段更易于直接测量。 276

限于相对论性时空的公理,只要求变化经典时空的不变性公理。相对于一个惯性系分别测量空间和时间是不变的。

我们现在把光速的不变的常数 c 附加于结构$(A, B, \mathcal{F})$。

定义 1. 一个结构 $\mathfrak{A} = (A, B, \mathcal{F}, c)$ 是具有惯性参照系的一个(四维的)限于相对论性的时空结构,当且仅当,满足下列公理:

公理 1. 与第 2 节的定义 1 的仿射公理相同。

公理 2. 与第 2 节的定义 1 的惯性系公理相同。

---

① 等式(1)是为了便于计算,但在概念上自然是除以 c,以便在任意给定的坐标系 f 中,固有时间的单位是时间单位,即

$$固有时间\ {}_f(x, y) = \frac{1}{c} I_f(x, y)。$$

公理 3.（光速的不变性）。正实数 c，代表在真空中测量光速，相对于 $\mathcal{F}$ 中的任何一个惯性系 f，都是不变的，即恒定的。

公理 4.（惯性线段的固有时间的不变性）。对于 A 中的任意两点 a 和 b，使得对于在 $\mathcal{F}$ 中的某个惯性系 f，$I_f^2(ab) > 0$，线段 ab 的固有时间在 $\mathcal{F}$ 中是不变的，即对于 $\mathcal{F}$ 中的任意两个惯性系 f 和 f'，

$$I_f(ab) = I_{f'}(ab) \tag{III}$$

公理 5.（不变的惯性系的闭合性）。如果 f 在 $\mathcal{F}$ 中，φ 是 (R^4, B_R) 的任意仿射变换，使得 f 和 $\varphi \circ f$ 满足(III)，那么 $\varphi \circ f$ 在 $\mathcal{F}$ 中。

我们下一步需要从形式上定义一个洛伦兹矩阵概念，然后，定义洛伦兹变换概念。

定义 2. 一个(4 阶)矩阵 $\mathbf{A}$ 是一个洛伦兹矩阵，当且仅当，存在实数 β、δ、一个三维向量 $\boldsymbol{u}$ 和一个 3 阶正交矩阵 $\boldsymbol{E}$，使得

$$\beta^2\left(1 - \frac{\boldsymbol{u}^2}{c^2}\right) = 1$$

$$\delta^2 = 1$$

$$\mathbf{A} = \begin{pmatrix} \boldsymbol{E} & \mathbf{0} \\ \mathbf{0} & \delta \end{pmatrix} \begin{pmatrix} \boldsymbol{I} + \dfrac{\beta - 1}{\boldsymbol{u}^2}\boldsymbol{u}^*\boldsymbol{u} & -\dfrac{\beta \boldsymbol{u}^*}{c^2} \\ -\beta \boldsymbol{u} & \beta \end{pmatrix}$$

(在这个定义中和其他地方，如果 $\mathbf{A}$ 是一个矩阵，那么，$\mathbf{A}^*$ 是它的转置矩阵，以及像 $\boldsymbol{u}$ 一样的向量都是向量——因此，$\boldsymbol{u}^*$ 是一个列向量。①)定义 2 中的各个量的物理解释应该都是显而易见的。数字 β 是洛伦兹收缩因子。当 $\delta = -1$ 时，我们有时间方向

① 为了简化符号，我有时省略列向量中的转置符号 *。

的反转。矩阵 $\boldsymbol{E}$ 表示 f 的空间坐标的一种旋转，或者，伴有偏转的旋转。向量 $\boldsymbol{u}$ 是两个参照系的相对速度。为了进一步参考，我们可能会注意到，每一个洛伦兹矩阵都是非奇异的。

定义 3. 一个洛伦兹变换是一个把 R_4 映射到自身的一一对应函数 φ，使得存在着一个洛伦兹矩阵 $\mathbf{A}$ 和相对于 $\mathcal{F}$ 中的一个参照系 f 的一个四维向量 $\boldsymbol{b}$，所以，对于 R_4 中的所有 $\boldsymbol{x}$，

277

$$\varphi(\boldsymbol{x}) = \boldsymbol{x}\mathbf{A} + \boldsymbol{b}$$

向量 $\boldsymbol{b}$ 的物理解释是清晰的。它的前三个坐标表示空间坐标原点的平移，它的最后一个坐标表示时间原点的平移。定义 3 解释了每一个洛伦兹变换都是 R_4 的一个非奇异仿射变换，即我们将在几种语境中用到的一个事实。

在不变性定理的证明中，方便的是，运用鲁宾和苏佩斯(1954)证明了的关于洛伦兹矩阵的一个预备定理，而且，这个预备定理只是一个直接计算的问题。

预备定理 19. 一个(4 阶)矩阵 $\mathbf{A}$ 是一个洛伦兹矩阵，当且仅当，

$$\mathbf{A}\begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & -c^2 \end{pmatrix}\mathbf{A}^* = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & -c^2 \end{pmatrix}$$

现在我们陈述不变性定理：

定理 1. 根据实数域上的仿射四维笛卡儿空间，惯性系的限于相对论性的时空结构的表征，在洛伦兹变换群下，是不变的。

证明： 设 f、f' 是 $\mathcal{F}$ 中的两个惯性系。像前面一样，对于 A 中的 a，$f(a) = \boldsymbol{x}$，$f_1(a) = x_1$，$f'(a) = \boldsymbol{x}'$，等等。我们考虑变换 φ，使得对于在 A 中的每个 a，$\varphi(\boldsymbol{x}) = \boldsymbol{x}'$。借助于公理 1 与公理 2 和定理 1.2，存在一个(4 阶)非奇异矩阵 $\mathbf{A}$ 和一个四维向量 $\boldsymbol{b}$，使得对于 R^4 中的每一个 $\boldsymbol{x}$，

$$\varphi(\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{A} + \boldsymbol{b}$$

这个证明简化为表明,$\boldsymbol{A}$ 是一个洛伦兹矩阵。

设

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{D} & \boldsymbol{e}^* \\ \boldsymbol{f} & g \end{pmatrix} \tag{3}$$

并且设 α 是(在 f 中)一条光线,使得对于 α 的任何两个不同的点 $\boldsymbol{x}$ 和 $\boldsymbol{y}$,如果 $\boldsymbol{x} = (Z_1, t_1)$ 和 $\boldsymbol{y} = (Z_2, t_2)$,那么

$$\frac{Z_1 - Z_2}{t_1 - t_2} = \boldsymbol{W} \tag{4}$$

显然,根据公理 3,$|\boldsymbol{W}| = c$。现在设

$$\boldsymbol{W}' = \frac{Z'_1 - Z'_2}{t'_1 - t'_2} \tag{5}$$

从(3)、(4)和(5),我们有

$$\boldsymbol{W}' = \frac{(Z_1 - Z_2)\boldsymbol{D} + (t_1 - t_2)\boldsymbol{f}}{(Z_1 - Z_2)\boldsymbol{e}^* + (t_1 - t_2)g} \tag{6}$$

(6)式右边的所有项除以 $t_1 - t_2$ 并运用(4),我们得到

$$\boldsymbol{W}' = \frac{\boldsymbol{W}\boldsymbol{D} + \boldsymbol{f}}{\boldsymbol{W}\boldsymbol{e}^* + g} \tag{7}$$

再一次从公理 3 得到

$$|\boldsymbol{W}'| = c \tag{8}$$

既然 $|\boldsymbol{W}'| = c$,所以,我们通过(7)的平方有　　278

$$\frac{\boldsymbol{W}\boldsymbol{D}\boldsymbol{D}^*\boldsymbol{W}^* + 2\boldsymbol{W}\boldsymbol{D}\boldsymbol{f}^* + |\boldsymbol{f}|^2}{(\boldsymbol{W}\boldsymbol{e}^* + g)^2} = c^2 \tag{9}$$

因而

$$\boldsymbol{W}(\boldsymbol{D}\boldsymbol{D}^{*}-c^{2}\boldsymbol{e}^{*}\boldsymbol{e})\boldsymbol{W}^{*}+2\boldsymbol{W}(\boldsymbol{D}\boldsymbol{f}^{*}-c^{2}\boldsymbol{e}^{*}g)+|\boldsymbol{f}|^{2}-c^{2}g^{2}=0 \tag{10}$$

既然(10)对于任意一条光线都成立,所以,我们可以用$-\boldsymbol{W}$代替$\boldsymbol{W}$,并再次得到(10)。我们于是推出

$$\boldsymbol{W}(\boldsymbol{D}\boldsymbol{f}^{*}-c^{2}\boldsymbol{e}^{*}g)=0$$

但$\boldsymbol{W}$的方向是任意的,据此,

$$\boldsymbol{D}\boldsymbol{f}^{*}-c^{2}\boldsymbol{e}^{*}g=0 \tag{11}$$

现在设$\boldsymbol{x}=(0,0,0,0)$,$\boldsymbol{y}=(0,0,0,1)$,于是

$$I_{f}^{2}(\boldsymbol{x}\boldsymbol{y})=c^{2}$$

但从(3)很容易看出

$$I_{f'}^{2}(\boldsymbol{x}\boldsymbol{y})=c^{2}g^{2}-|\boldsymbol{f}|^{2}$$

因此,根据我们的基本的不变性公理(公理 4),

$$c^{2}g^{2}-|\boldsymbol{f}|^{2}=c^{2} \tag{12}$$

从(10)、(11)、(12)和事实$|\boldsymbol{W}|^{2}=c^{2}$,我们推出

$$\boldsymbol{W}(\boldsymbol{D}\boldsymbol{D}^{*}-c^{2}\boldsymbol{e}^{*}\boldsymbol{e})\boldsymbol{W}^{*}=|\boldsymbol{W}|^{2}$$

而且,由于$\boldsymbol{W}$的方向是任意的,我们断定

$$\boldsymbol{D}\boldsymbol{D}^{*}-c^{2}\boldsymbol{e}^{*}\boldsymbol{e}=\boldsymbol{I} \tag{13}$$

这里$\boldsymbol{I}$是单位矩阵。

现在通过基于(3)的直接计算

$$\boldsymbol{A}\begin{pmatrix}\boldsymbol{I} & 0\\ 0 & -c^{2}\end{pmatrix}\boldsymbol{A}^{*}=\begin{pmatrix}\boldsymbol{D}\boldsymbol{D}^{*}-c^{2}\boldsymbol{e}^{*}\boldsymbol{e} & \boldsymbol{D}\boldsymbol{f}^{*}-c^{2}\boldsymbol{e}^{*}g\\ (\boldsymbol{D}\boldsymbol{f}^{*}-c^{2}\boldsymbol{e}^{*}g)^{*} & \boldsymbol{f}\boldsymbol{f}^{*}-c^{2}g^{2}\end{pmatrix} \tag{14}$$

从(11)、(12)、(13)和(14),我们最终得到结果

$$\mathbf{A}\begin{pmatrix} \boldsymbol{I} & 0 \\ 0 & -c^2 \end{pmatrix}\mathbf{A}^* = \begin{pmatrix} \boldsymbol{I} & 0 \\ 0 & -c^2 \end{pmatrix}$$

因此，依据预备定理 19，$\mathbf{A}$ 是一个洛伦兹矩阵。

证毕。

历史上的评论。正如已经陈述的那样，与经典物理学相应的情形完全相反，在狭义相对论的最初发展中，明显的不变性问题是很突出的。

爱因斯坦的前辈是沃伊特(Voigt)、斐兹杰惹(FitzGerald)、洛伦兹(Lorentz)以及彭加勒(Poincaré)，他们中没有一个人明确地阐释过相对论性的不变性原理。事实上，大多数努力的潜在动机仍然是为了寻找作为惟一首选参照系基础的绝对静止的以太。沃伊特(1887)显然实际上第一个公布了洛伦兹变换的一种看法，但是，他所关注的是多普勒频移。[1] 爱尔兰物理学家斐兹杰惹深受著名的迈克耳孙—莫雷实验的影响，这个实验得出了地球在以太中运动的结果为零，他第一个提出相对于测量参照系以速度 v 运动的一个物体的长度在其运动方向上的收缩因子是 $\sqrt{1-v^2/c^2}$。但是，对于斐兹杰惹来说，这种变化是真实的，正如他指出的那样，“……当物体穿过以太时，他们的长度发生了改变……”洛伦兹 (1892)在不知道斐兹杰惹文章的前提下，也提出了这样一个收缩因子。然后，洛伦兹(1895)根据作为他的电磁理论发展的一个组成部分的地面上的光学实验，推导出了一个有用的洛伦兹变换形式。彭加勒(1898)非常清晰地讨论了判断远距离的两个事件同时发生的明显的操作困难，但他只是作出了纲领性的评论。

279

---

① 这里和在爱因斯坦的出版物(1905/1923)之前，我引用了派斯(1982)的出色的历史分析。

爱因斯坦的出版物(1905/1923)是决定性的事件,对不变性作出了明确的陈述,这是从麦克斯韦电磁学理论到运动学的狭义相对论的至关重要的一步。下面是这篇著名论文的开头两段的英文翻译(1923)。

> 众所周知,麦克斯韦的电动力学——正如现在通常所理解的那样——当被应用于运动物体时,导致了一些不对称,而这些不对称似乎不是现象固有的。以一块磁铁与一个导体的相互电磁作用为例。这里可观察的现象只依赖于导体和磁铁的相对运动,而习惯的观点认为,在这两个物体中,是这个物体在运动,还是那个物体在运动,是截然不同的两回事。因为如果是磁铁运动,而导体静止,那么在磁铁周围会形成一个具有一定能量的电场,位于电场中的导体部分会产生一种电流。但是,如果磁铁静止,而导体运动,在磁铁周围就不会产生电场。然而,在这个运动的导体中,我们发现了一个电动势,这种电动势本身并没有相应的能量,但是,它也会引起——假定在这里讨论的两种情况下相对运动是一样的——同前一种情况通过电力产生的那些路径和强度一样的电流。
>
> 诸如此类的例子,与企图发现地球相对于"光媒质"的任何一种运动的失败一起,间接地表明,电动力学现象和力学现象没有与绝对静止的思想相对应的特征。它们反而使人联想到,正如少数一阶量所表明的那样,同样的电动力学定律和光学定律,对力学方程很好地成立的一切参照系,将是有效的。① 我们将要把这个推测(此后,这个要旨被称为

---

① 这里的参照系是指经典物理学的惯性系,当然,这些惯性系同样适用于狭义相对论。

"相对性原理")提升到一个公设的地位,并且,还要引入另一个公设,这个公设只在表面上与前一个公设相矛盾,前一个公设是,光在真空中总是以一个确定的速度传播,与放射体的运动状态无关。这两条公设以静止物体的麦克斯韦理论为基础,足以得到一个简单而一致的运动物体的电动力学理论。"光以太"的引入将证明是多余的,因为这里提出的观点,将不需要一个提供特殊性质的"绝对静止的空间",也不需要把一个速度矢量分派给产生了电磁场过程的真空的一点。

(Einstein, 1905/1923: pp. 37 - 38)

爱因斯坦文章的第一段的重点是,定义远距离事件的同时性,彭加勒(1898)曾较早地暗示过,但只在这里的陈述中,才明确而大胆地用了同时性的定义。我给出一个等价的仿射几何表述。设 a 是任意一个时空点,即仿射点。设 L 是通过 a 的一条惯性线;这里 L 作为定义同时性的惯性系的时间轴。如果两条线成为一个光平行四边形的两条对角线,即仿射平行四边形的对角线,像前面定义的那样,这个仿射平行四边形的边是光线的线段,那么任何一条仿射线都正交于 L。 280

于是,

$$\mathcal{S}(a, L) = \{b: \text{直线 } ab \text{ 正交于 } L\} \cup \{a\}$$

是相对于 L 的与 a 同时的点集。注意,同时性的定义必须是相对于一个给定参照系的。经典物理学的绝对同时性被放弃了。

在对狭义相对论的几何结构进行定性的公理化时,自然会假设,对于每一个仿射时空点 a 和通过 a 的任何一条惯性线,集合 $\mathcal{S}(a, L)$ 是相对于仿射的介中性关系和满足标准的欧几里得全等公理的全等关系 $\approx$ 的一个三维欧几里得空间。

爱因斯坦在1905年的这篇著名文章的第二部分，明确地陈述了他所用的两个原理：

> 下列反思是建立在相对性原理和光速不变性原理的基础上的。我们把这两个原理定义如下：
>
> 1. 经历变化的物理系统的态所遵守的定律不会改变，无论这些态的变化是参照两个匀速平移运动的坐标系中的哪一个坐标系。
>
> 2. 任何一条光线都以确定的速度 c 在"静止的"坐标系中运动，不管这条光线是由静止物体发出的，还是由运动的物体发出的。
>
> (Einstein 1905/1923, p. 41)

然后，用这些原理表明时间和长度的相对性。

在第3部分推导出一种洛伦兹变换形式，但是，正如前面注意的那样；并不知道洛伦兹较早文章。第4部分考虑了所推出的这种变换的物理意义。这里，爱因斯坦引入了一个运动物体的固有时间 τ，即"相对于运动系统是静止的"时间，并推出等式

$$\tau = t\sqrt{1-\frac{v^2}{c^2}}$$

在第5部分中，推出了相对论性的速度合成公式。

这篇文章的其余部分是介绍电磁学理论，特别是推出了麦克斯韦方程在洛伦兹变换条件下的协变性。①

爱因斯坦于1905年6月30日把他的论文提交给《物理学年刊》。彭加勒的文章则是于1905年5月提交的，而且，他的第

① 在4.4节已经分析了协变性和不变性的关系。

二篇文章是在 1905 年 7 月完成的，在 1906 年发表。当然，他们两个人当时都不知道对方发表的文章。彭加勒在第二篇论文中对麦克斯韦方程的协变性进行了很详细的证明。但是，非常令人奇怪的是，彭加勒显然完全没有抓住狭义相对论的意义。后来 1909 年在哥根廷的讲座（这些讲座于 1910 年出版）中，他仍然超出爱因斯坦在 1905 年文章的第二部分所用的和上面引证的这两个原理的范围，引入了斐兹杰惹-洛伦兹收缩的物理实在作为第三个假设，这显然混淆了狭义相对论的本性。 281

后来的定性公理进路。在狭义相对论的情形中，追寻时序的基本二元关系是整个几何学的一个充分的结构基础。许多年前，罗布已经证明了这一点（1911，1914，1921，1928，1930，1936）。对他的思想的最广泛的阐明，能在他的 1936 年的书中找到。对罗布工作的详细讨论和某些推广能在温妮（Winnie）的著作中（1977）找到。亚历山德罗夫（Alexandrov，1950，1975）表明，罗布的惟一二元关系的数值表征的不变性足以蕴含着洛伦兹变换。后来，塞曼（Zeeman，1964，1967）也独立地表明了这一点。令人惊奇的是，罗布根本没有讨论这个自然的问题。

自从罗布的早期著作出版之后，关于狭义相对论几何的定性公理进路，已经发表了许多文章。较早的进路是威尔孙（Wilson）和刘易斯（Lewis）的进路（1912），他们给出了强调平行变换的仿射二维情形的不完备的公理化。赖兴巴赫（Reichenbach，1924）给出了在哲学上令人感兴趣的和在直观上吸引人的一种公理讨论，但是，这种讨论更多是在物理学精神的鼓舞下进行的，而不是在一种清楚而精确的几何公理化精神的鼓舞下进行的。他强调了对同时性和约定主义的哲学问题的分

析。沃克(Walker,1948,1959),建立在英国物理学家米尔恩(E. A. Milne,1848)的相对论的物理学进路的基础上,由于把一个质点区分为质点—观察者,所以,把除时空点之外的质点看成与这些点的某些类(直观上,这些类是质点的路径)和罗布追寻的二元关系一样是未定义的。此外,他假设了任意两个质点的时空点之间信号对应的一种基本关系。也许,沃克进路的最不能令人满意的特征是,信号映射是复函数,这些函数做了需要通过认真地提出更基本的和更可能直观地可以描述的运算或关系才可以做到的事情。赫金(Hudgin,1973)提供了与沃克的进路既独立又相关的另一条进路,他首先从类时测量开始,并用一个通信函数表示信号的对应。同沃克一样,赫金运用了米尔恩的物理学思想。

德默特尔(1972)发表了对罗布的详细讨论,并且通过给出一种向量空间描述,提供了关于罗布进路的另一种观点。拉策(Latzer ,1972)表明,有可能把光信号的一种二元关系看成是基元。他没有给出一种新的公理化分析,但表明如何能够根据他的光的对称关系来定义罗布的排序关系。从一种纯几何学的观点来看,有一个充分的论据支持采纳这样一种对称关系,因为根据纯几何学的理由,区分时间方向的方式完全是随意的。实际上,从亚历山德罗夫更早的文章(1950,1975)中得出的结论是,拉策的二元基元是适当的。

282 大概与罗布的工作相比,近来最详尽的努力之一是舒茨(Schutz)的工作(1973),舒茨受到了沃克工作的影响。舒茨的公理系统在两方面超越了沃克的公理系统。第一,沃克陈述的公理还没有强到足以完备地描述限于相对论的时空;第二,在这些公理中,多于一半的公理具有的形式,与沃克给出的那些形式没有密切的关系,虽然在几种情况下,根据物理学的直

觉，它们有相当强的感染力。正如舒茨表明的那样，他的著作的最直接重要的前期成果是塞凯赖什(Szekeres，1968)的文章，塞凯赖什的进路类似于沃克的进路，但塞凯赖什把质点和光信号都看成是对象，而不是把光信号看成在形式上具有由信号的二元关系所描述的特征。后来，舒茨(1979)简化了他的公理，并使这些公理具有更好的形式，但本质上依据了同样的进路。

蒙迪(Mundy)最近的工作(1986a，1986b)也是重要的。在蒙迪的文章(1986b)中，他对明可夫斯基几何的物理学内容进行了详细的分析。他也包括了对更定向于物理学的其他公理化的概述，比如，梅尔伯格(Mehlberg，1935，1937)的公理化。蒙迪在他的公理化中把光速的不变性强调为是明可夫斯基几何的物理学核心。

在蒙迪的文章(1986a)中，他给出了一种光学的公理化，这种公理化接近于罗布公理化的精神，尽管惟一的二元基元接近于拉策的思想。蒙迪的公式化肯定是对罗布的公式化的一种改进，并且是迄今为止依赖于惟一的二元关系的最好的公理集。

在更新近的著作中，戈德布拉特(Goldblatt，1987)也许是在文献中第一次强调，从相对论的时空概括出的正交性概念，能够发挥的核心作用。在具有标准内积的欧几里得的向量空间中，两个向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是正交的，当且仅当，$\boldsymbol{x}\cdot\boldsymbol{y}=0$。我们运用不定的二次度量 $I^2(\boldsymbol{x},\ \boldsymbol{y})$，得到了下列观点：当两条直线相正交时，一条直线一定是类时的(一条惯性线)，另一条直线一定是类空的(一条分隔线)；一条光线正交于自身。按照公理，戈德布拉特为基于介中性的仿射公理，增加了基于 ab 正交于 cd 的四元关系的一些附加的公理。

6.4 如何确定视觉空间是否是欧几里得空间[①]

过去的哲学家们一直声称,视觉空间是否是欧几里得空间这个问题的答案能够通过先天的或纯哲学的方法找到。今天,大概只有极少数著名的康德主义者持有这样一种观点。人们普遍认为,各种方式的回答必定都是经验的,但可能因为间接的原因是经验的。通过物理空间是欧几里得空间的物理学论证,然后通过必要的视觉空间一定是欧几里得空间的感知的概念论证,能够确定这一点。在某种程度上,这一定是许多外行的观点:他们在很高的近似程度上承认,物理空间是欧几里得空间,因此,自动持有视觉空间是欧几里得空间的观点。

我首先从这样的问题开始:我们如何检验视觉空间是欧几里得空间这个命题呢?本节关注方法论问题,并包括对与视觉现象相关的几何层次结构的简要概述。下一节评论为这个问题提供的许多答案。我既考察了哲学的断言,也考察了心理学的断言。第 7 节关注由现有的各种答案产生的核心问题。

283

在许多方面,看起来会是这个问题的最自然的数学进路,也一直是实验上用得最多的方法。它是由考虑一个有穷点集组成的。在实验上,这些点接近于暗室里显示出的低亮度的小的点光源。这种框架的直观思想是,只使有穷数目的点光源成为看得见的,而且,使这些光源的强度非常之低,足以排除周围的亮光。第二步是要求人们作出视觉判断,陈述特定的几何关系在这些点之间是否成立。例如,点 a 和点 b 之间的距离等同于点 c

① 本节和下一节的许多资料都来源于苏佩斯(Suppes,1977a)。

和点 d 之间的距离吗？（此后，在这种讨论中，我所指的点，应该理解为是，我想到的根据点光源的这种物理实现。）另一类问题可能是，点 a、b、c 形成的角，在测量时，全等于或等同于点 d、e、f 形成的角吗？

这些判断的另一条进路，不是质问已知的点是否有特定的关系，而是允许作出判断的人利用其中的某些点。例如，先固定点 a、b 和 c，然后，要求他调节 d，使得 c 和 d 之间的距离等于 a 和 b 之间的距离。尽管我对这些问题的表述听起来好像它们在特征上是度量的，但是，它们通常是定性的性质——例如，我把线段的全等表述为具有相同的距离。我们没有把度量的要求强加于作出这种判断的那些人。例如，在与我们的问题相关的实验中，没有人会自然地要求把两点之间的距离大约规定为 1.3 米，或者比如说，把一个角确定为 21 度。

无论是基于固定关系，还是通过调节点的位置，一旦得出这些判断，要问的形式的或数学的问题是，在二维或三维欧几里得空间中，是否嵌入了这种有穷关系结构。维数取决于实验的特征。在许多情况下，把这些点限于一个平面，因此，要求在二维空间中嵌入这些点；在其他情况下，把这些点嵌入三维空间是适当的。我说的**有穷关系结构**的意思是指，它的定义域是有穷的这样一个关系结构。举一个简单的例子来说，假设 A 是这个有穷点集，我们要求的判断是，对这些点之间的等距作出判断。设 E 是等距的四元关系。于是，说能够把有穷关系结构 $\mathfrak{A}=(A, E)$ 嵌入三维欧几里得空间，就是说，存在着在 A 上定义的一个函数 φ，使得 φ 把 A 映射到实数的三元组的集合，并且使得对于 A 中的每一个 a、b、c 和 d，满足下列关系：

$$ab\ E\ cd \text{ 当且仅当} \sum_{i=1}^{3}(\varphi_i(a)-\varphi_i(b))^2=\sum_{i=1}^{3}(\varphi_i(c)-\varphi_i(d))^2,$$

284 这里 $\varphi_i(a)$是 $\varphi(a)$的第 i 个坐标。注意,映射到实数的三元组,恰好是把视觉点映射到三维欧几里得空间的一个笛卡儿表征。

在原则上,回答由这个嵌入程序产生的问题是简单的。因此,已知一个人对点之间的等距作出视觉判断的数据集合,我们能以一种明确的和作图的数学方法确定,这种同构嵌入是否可能。

然而,立即产生了一个问题。这个问题能够通过考虑类似的物理学情境来把握。假设我们正在观测星体,并且想要检验一个相似的命题,或者,某个更复杂的天体力学命题。我们面临着在天文学史和测地学史上早已公认的问题:数据不一定精确地与理论模型相符合。提出这种观点的经典方式是,产生了测量误差,然后,我们的问题是确定,在测量误差的范围内模型是否与数据相符合。在考察水星近日点进动的数据时(这是爱因斯坦的广义相对论的重要检验之一),数据分析的最繁琐和最困难的问题是确定,在估计的测量误差范围内理论和观测是否一致。

例如,拉普拉斯(1799/1966)空前成功地运用了这样的方法。他根据太阳系的一些特殊现象,例如,木星和土星运动中的不规则性考察数据,然后,提出的问题是,这些观测到的不规则性是由测量误差引起的,还是由存在的"不变的"原因引起的。当这种不规则性太明显,以至于不能根据测量误差加以解释时,他就寻找一种不变的原因来说明为什么会背离这种较简单的现象模型。在所提到的木星和土星运动中不规则性的情况下,他能把这些不规则性解释为是由两个天体的相互的引力造成的,在它们运动的简单理论中一直忽略了这一点。但是,拉普拉斯的情境在下面的重要方面却与现在的情境不同。他考察的数据已经是以定量的形式给出的,而且,根本没有数值表征问题。我

们的问题是,我们从定性判断出发,就面临着同时选定一种测量和确定其测量误差的问题。在现在的情况下,因为关于测量误差的统计问题是复杂的和微妙的,所以,为了简单起见,我将忽略这些问题,但是,在对实验数据的任何详细的分析中必须研究这些问题,承认这一点是绝对必要的。

回到把有穷点集中的定性关系嵌入到一个已知空间的形式问题,令人惊讶的是发现,在当前语境中需要的这种结果,实际上,在几何的数学文献中却不会出现。存在着大量关于有穷几何的文献;例如,在登博夫斯基(Dembowski)的著作(1968)中列出了一千二百多个参考文献。此外,考虑有穷几何的传统至少可追溯到20世纪初。维布伦等人的这些几何作图法是证明公理等的独立性的模型的一个丰富来源。① 另一方面,在登博夫斯基的权威调查中,引用最多的文献几乎全是具有相对弱结构的投影几何和仿射几何。从数学的观点来看,这些结构一直是相当令人感兴趣的,与抽象代数中的不同问题联系在一起。相应的一种较强类型的有穷几何的理论,例如,有穷欧几里得几何、有穷椭圆几何或有穷双曲几何,很少有所发展。结果,实验文献并没有直接研究这样的有穷几何,尽管它们一方面是更弱的有穷几何的一种自然推广,另一方面是有穷测量结构的一种自然推广。(更详细的讨论参见第8节。)

285

视觉空间的几何特征的第二条基本的方法论进路是假定,已经存在一个标准的度量表征,然后,考察哪种空间与数据最相符。在福莱(1964,1972)的不同出版物中,会找到这种方法论的一个极好的例子。福莱根据实验表明,大小距离不变性假说(size-distance invariance hypothesis),即断言感知大小距离的

① 参见维布伦(1904)以及维布伦和杨(Young,1910,1918)的著作。

比率等于物理的大小距离的比率,是非常错误的。同时,他也表明,感知视角比物理角度约大百分之十。进行这些研究根据的假设是,满足各种各样的更原始和更基本的公理。相反,吕内堡(Luneburg,1948)假设,感知视角等于物理角度,也就是说,这两者之间的变换是保角变换,但运用这个假设的基础是下列假设的所有变种:物理空间和视觉空间都是常曲率的齐性空间,即都是黎曼空间,并且在本质上,吕内堡并不打算以任何严格的方式检验具有常曲率的齐性空间的非常丰富的假设蕴含的许多结果。换言之,在第二条进路中,没有认真努力地通过检验表明,一个已知的空间类型成立的所有公理是否都会得到满足。

第三条进路是回到关于空间本性的著名的亥姆霍兹-李(Helmholtz - Lie)问题,并用连续性和运动的问题取代有穷性。黎曼(Reimann,1866 - 1867)在 1854 年著名的讲座中讨论了几何基础依赖的假设。十多年之后,亥姆霍兹(1868)在一篇题为"关于基本几何事实(Über die Tatsachen, die der Geometrie zu Grunde liegen)"的论文中作了回应。亥姆霍兹论文的基本论证是,虽然任意的黎曼空间是可想到的,但是,作为一个基本特征,现实的物理空间具有刚体的自由运动性。从数学的观点来看,这样的运动在度量几何中被描述为把一个空间变换到自身保持距离不变。这样的变换被称为**等距映射**,意指在度量上是同构的。由于这个主题得到了广泛的数学发展(关于当代的评论,参见 Busemann,1955,48 节,或 Freudenthal,1965),大量极好的形式结果有利于促使把这种观点运用到视觉空间的特征研究中。在亥姆霍兹—李进路的各种公理化的条件下,可以证明,惟一满足这些公理的空间是下列三种基本空间:欧几里得空间、双曲空间和椭圆空间。

286 从哲学的观点来看,重要的是承认在分析视觉空间的本质

时，考虑连续性和运动可能比有穷空间的数学上更初步的特性更基本。不幸的是，我不能转述用亥姆霍兹-李进路作为研究视觉空间本性的方法的任何实验文献。尽管这已经隐含在下面报告的一些结果中，但是，很难把实验结果解释为是满足一个自由运动的公理。让我更清楚地阐明这种观点。一些实验研究得出这样的结果：在刚才定义的意义上，视觉空间不可能是基本的，但是，这些研究显然没有运用由在回应亥姆霍兹-李问题（即对允许刚体的自由运动的空间进行研究，）时相伴随的丰富的数学进展所建议的那种运动进路。

第四条进路是根据图像文法和知觉情景分析的进路，这不包括在本章考虑的主要文献当中。这条进路的不断增加的文献来源于对以设计计算机程序和具有基本感知能力的外围设备为中心的模式识别问题的特别回应。尽管这条进路具有的形式特征完全不同于所考虑的其他进路，也一直没有被用来直接提出关于欧几里得的视觉空间特征的问题，但是，应该提到的是，因为它提供了在许多方面从心理学的角度看都很自然的一条进路，并且，它在某些方面比分析视觉空间时迄今所用的大多数经典几何进路更密切地与知觉心理学相联系。[关于较早文献的初步介绍和参考文献可以在苏佩斯和罗特迈尔(Rottmayer)的文章(1974)中找到；福(Fu)给出了一种百科全书式的评论(1974)。最近二十五年的文献太多了，不可能只引用一种文献概述。近来重要的统计趋势在阿蒂亚斯(Attias,1999)、布赖曼(Breiman,1999)以及弗里德曼(Friedman)、黑斯蒂(Hastie)和蒂布希拉尼(Tibshirani)的文章(2000)中有所讨论。]

一种典型的图像文法具有下列特征。把特定长度和特定方向的几条有穷线段或有穷曲线连接在一起，作为形成更复杂的几何图形的基本要素。在模式识别的文献中一个典型问题是提

供这样一种连接(不一定是一维的),以便构成手写字符,或者,作为已经很关注的一个特例,识别手写的数学符号。这些进路通常被贴上图像文法的标签,因为他们采纳在写短语结构文法的数学语言中所用的这条进路,产生语言表达。图像文法事实上能够被描述为是语境无关的(context free)、语境敏感的(context sensitive),等等,即依赖于产生规则的准确字符。遗漏掉的问题是,由图像文法产生的图形集能够被嵌入欧几里得空间或一个基本字符的其他度量空间吗?从知觉理论的观点来看,这个问题似乎有某种概念兴趣。显然不是与模式识别理论一样重要。把知觉建立在原始概念基础上的图像文法,似乎比经典几何中熟悉的更抽象的概念更自然得多。此外,至少在某287 种程度上,关于独特的脑神经元的行为的大量结果支持:许多动物的视觉系统中有专门的神经元检测图像文法中所用的这种明确特征。

几何的层次结构。宣布视觉空间不是欧几里得空间的那些人,通常已经想到了一种界定明确的替代选择。最流行的候选方案声称,视觉空间要么是椭圆空间,要么是双曲空间,尽管在一些实验工作中隐含了某些更偏激的论题。

各种不同的几何如何在层次结构的意义上相关,完全不是一个简单的问题,因为根据不同的特化方法,一种几何可以从另一种几何中获得。为了探讨视觉空间,一个合理的自然的层次结构如图 1 所示。在这个图中,我指的也是几何,而不是空间,尽管从一种特定的概念观点来看,后者更可取。我一直坚持用有关视觉空间的文献中的传统遵循几何的语言。这里考虑的最弱的几何要么是图中最上面一行左边的射影几何,要么是最上面一行右边的有序几何。射影几何存在着各种自然的原始概念。无论如何,基本概念是入射概念,而且,在第三章中,一旦引

入顺序,也就引入了分隔概念(the concept of separation)。相反,有序几何建立在时髦的欧几里得几何标准中对三个点成立的介中性的一种三元关系的基础之上,但是,只建立在介中性基础上的公理,当然比双曲几何或欧几里得几何所要求的那些公理更弱。如果不考虑技术细节,平面的椭圆几何可以通过下列方式根据射影几何得到:即把射影几何定义为在射影平面上留下一个不变的虚椭圆的射影直射变换群相对应的几何。尽管在考虑视觉空间时,椭圆几何一直是重要的,正如我们后面将看到的那样,椭圆几何的细节是复杂的和微妙的,而且,就我所知,实际上还没有针对严格的实验数据在细节上适当地加以研究。

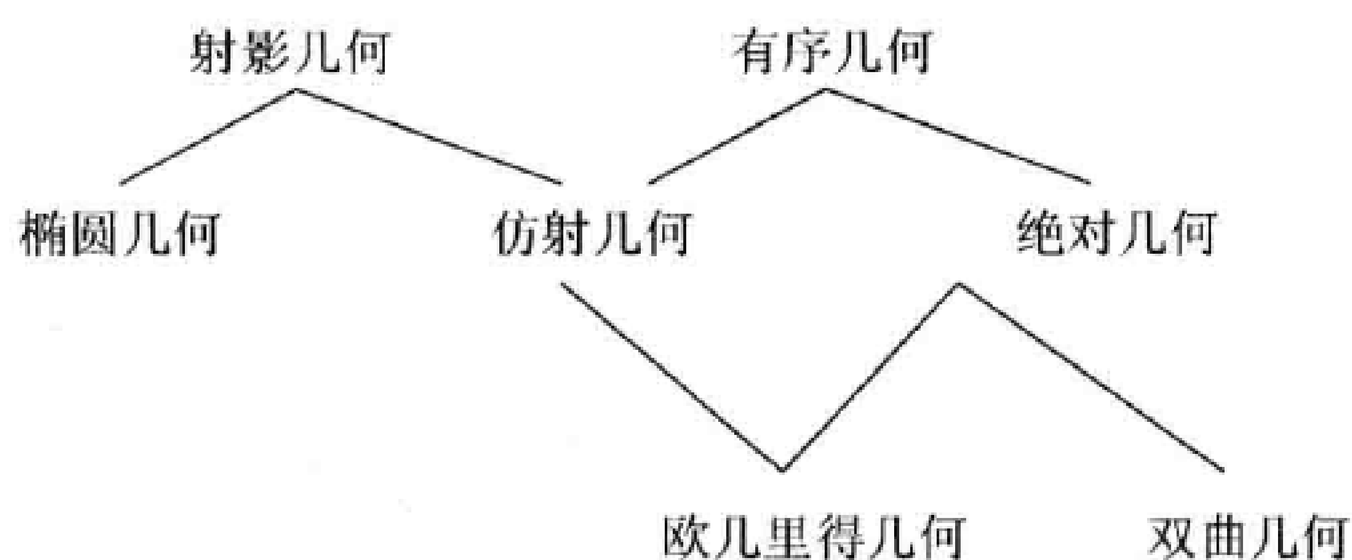


图 1　几何的层次结构

现在转向图 1 的右边,仿射几何可以通过增加欧几里得公理从有序几何中获得:即已知一条直线和这条直线之外的一点,最多只能画出一条直线,(i) 通过这个点,(ii) 在由这个点和这条线形成的平面上,以及(iii) 不与这条线相交。从图 1 中的有序几何的另一个分支,我们可以通过增加线段全等的概念得到绝对几何,这恰好是前面提到的等距概念。我们把欧几里得公理加到绝对几何中,可获得欧几里得几何,并且,我们把对欧几里得公理的否定加到绝对几何中,可获得双曲几何。这些是对绝对几何仅有的两种推广。已知与通常作出视觉空间要么是 288

欧几里得空间要么是双曲空间的断言相关的绝对几何的基本特征，有点令人感到惊讶的是，绝对几何的公理对视觉空间是否成立，还没有进行过详细的实验研究。

还有另一种方式来根据度量空间组织几何层次结构。回想一下，一个**度量空间**是一个对(A, d)，使得A是一个非空集，d是在笛卡儿积$A \times A$上定义的一个实值函数，使得对于A中的所有a、b、c，(i) $d(a, a) = 0$，并且，如果$a \neq b$，则$d(a, b) > 0$；(ii) $d(a, b) = d(b, a)$；以及(iii) $d(a, b) + d(b, c) \geqslant d(a, c)$。集合$A$的元素被称为**点**。第一个公理断言，距离是正的，除了同一点间的距离等于零之外。第二个公理断言，距离是对称的；即它只是无序点对的一个函数，而不是它们的顺序的一个函数。第三个公理是三角不等式。

对于知觉理论重要的大多数度量空间都有这样的特性：通过一条线段可以把任意两点连接起来。这样的空间被称为含有增加线段的度量空间。我们能够把这样的空间自然地划分为两个主要子集：仿射度量空间和与坐标无关的度量空间。根据所谓的明可夫斯基度量，通过对这些子集中的每个子集的进一步具体化，我们很自然地通向欧几里得空间、双曲空间和球面空间，也通向对欧几里得度量的推广。与坐标无关的度量的一个重要子集是黎曼度量族。可以表明，既是黎曼度量空间又是仿射度量空间的惟一空间，要么是欧几里得空间，要么是双曲空间。我们将不详细地运用这些概念，但重要的是提到运用度量空间的这种替代的层次结构，与运用图1中显示出的更经典的层次结构一样自然。我在对几何层次结构的这种简要概述中引入的所有概念，在几何的数学文献中都是常见的。更多的细节可以在苏佩斯等人的著作(1989，第12章和13章)中找到。

6.5 视觉空间的本性：实验的回答与哲学的回答

我在这一节的主要目的是对视觉空间本性的现有回答进行一种概述。自然要以欧几里得的《光学》为出发点，这本书是现存最早的关于数学光学的著作。正如在 2.4 节所陈述的那样，重要的是强调指出，欧几里得的《光学》实际上是一种视觉理论，不是关于物理光学的著作。

在这种概述中，我们会反复遇到欧几里得限于的单眼视觉。然而，我们应该注意到，他证明了包括双眼视觉在内的几个命题；例如，"如果眼距大于球的直径，那么，会看到多半个球"。欧几里得不只限于一些简单的几何光学，而且确实关注视觉理论，从下列命题来看，这一点是显而易见的："如果一个圆弧与眼睛位于同一平面，那么这个圆弧看起来是一条直线。"这种命题是后来的那些理论的前身——例如，托马斯·里德（Thomas Reid）的理论（1764/1967）——它们强调了视觉空间的非欧几里得的特征。

我很快略过从欧几里得之后到 18 世纪这段时期，不是因为这么长的中间时期没有令人感兴趣的问题，而是因为关于视觉空间特征的看法，似乎没有发生显著的变化，或者至少，倘若发生了某些变化，我也不熟悉。例如，最近，我看到了大卫·林德伯格（David C. Lindberg，1970）翻译的 13 世纪的约翰·佩尚姆（John Pecham）的著作《数学评论》（*Perspectiva Communis*），尽管著作本身和林德伯格对它的评论充满了关于光学的其他很重要的有趣问题，例如，光的因果性理论，但我觉得在当前的语境中没有什么可汇报的。 289

牛顿的《光学》(1704/1931)与欧几里得的《光学》有天壤之别。初始定义直到公理8才提到眼睛,而且当时是以很狭义的方式提到的。牛顿光学的命题几乎毫无例外地关注光的几何属性,特别是光的物理属性。只是在最后的几个质问中才有一些关于眼睛机制的任何推测,并且这些推测与现在的主题无关。

在牛顿的《光学》第一版出版五年后,1709年出版了贝克莱(Berkeley)的《视觉新论》(1709/1901)。贝克莱基本上以否定的方式讨论视觉空间几何问题。他提出的观点是,距离不可能被直接看到,并且,事实上,似乎把对距离的感知归类为一个触觉问题,而不是视觉(感觉)问题,因为眼睛肌肉的会聚在特征上是触觉的。他强调的观点是,在几何意义上,我们不能观察或计算由遥远的一点为顶点其边指向双眼的两个中心所产生的视角。他在讨论视角的感知时指出,"因此,既然那些角和线本身都不可能通过视力来感知,由此得出结论……大脑不是根据它们来判断对象的距离"(#13)。他在谈到距离时所说的话,也在谈到凭视力不可能直接感知的大小时说了。在(#53)这一段,他尤其否认用视觉世界的几何作为视觉感知基础的企图。

从这几段和别的段落明显地看出,对于贝克莱来说,视觉空间不是欧几里得空间,因为根本不存在对距离或大小的正确感知;至少,视觉空间不是三维的欧几里得空间。至于人们是否应该认为视觉空间起码是二维的欧几里得空间,他似乎也没有给出明确的阐述。我自己的倾向是断定,他对这个问题的观点是否定大于肯定。也许,一个充分的否定论证可能是由他的下列坚持构成的:即他坚持存在着最小可视度。正如他所提出的那样,"感觉的范围不是无限可分的这点是确定无疑的。存在着最小有形物和最小可视度;超出这个范围,感官就无法感知。大家的经验告诉了他这一点"(#54)。

事实上，直到这篇论文的结尾，贝克莱才解释清楚，二维几何也不是视觉空间或像我们所说的视觉域的一个特有部分。正如他在这篇论文的最后一段所言，“至此，我假设，显然，抽象的或视觉的广延都不会成为几何对象”。

于 1764 年首版的里德的《探索人类心灵》(1764/1967)是最令人感兴趣的。第 6 章研究视觉，第 9 节是很有名的一节，标题为“视觉几何”。有时，据说这一节正是非欧几何的前身，但倘若 290 这样，也必须把它看成是潜在的前身，因为里德作为视觉几何来明确讨论的几何，完全是根据球面几何阐述的，球面几何当然自古以来一直被看成是几何的一个固有部分。里德提出的观点显然在这一节的开头就得到了明确的阐述：“假设一只眼睛位于一个球的中心，对于这只眼睛来说，这个球的每一个大圆都有相同的外表，好像它是一条直线；因为眼睛觉察不到直接面向眼睛的圆的弧度。由于同样的理由，在这个球的一个大圆的平面上画出任何一条线，实际上无论曲直，都将被这只眼睛看见。”重要的是注意到，里德的视觉几何是单眼视觉几何。他在其他地方提到过双眼视觉，不过，几何的详细发展限于单眼几何。贝克莱和里德之间的重要对比是，里德以直接的、非形式的数学方式详细地发展了单眼几何，如此类似的发展在贝克莱那里是没有的。

丹尼尔斯(Daniels，1972)强有力地认为，里德的视觉几何不仅使用了球面几何，而且是对里德的二重椭圆空间的一种介绍。安格尔(Angell，1974)也作出了类似的论证。我同情这些论证，但是，似乎对我来说，它们走得太远，理由是相当简单的，丹尼尔斯或安格尔没有讨论。让我们回想一下 19 世纪末费利克斯·克莱因是如何创立椭圆几何的。他承认，非常类似于欧几里得几何或双曲几何的自然几何能通过把对径点(antipodal)看成是一个单点(a single point)从球面几何中获得。球面几何

作为一种几何,其发展与欧几里得几何的发展几乎是并列的,球面几何的困难在于,对应于两条线的两个大圆有两个交点,不是一个交点。然而,通过把两个对径点确认为一个单点,标准的欧几里得的很多公设仍然有效。很显然,里德没有如此确认对径点,因为他在第五命题中很明确地指出,"任意延长两条直线将会相交于两点,并且彼此交叉。"正是相交于两点而不是一点的这种特性,防止他的视觉几何完全成为椭圆几何,并迫使我们继续根据里德自己直接运用的球状模型思考视觉几何。

尽管亥姆霍兹的广泛的经验工作和理论工作都是关于视觉的,但是,他没有更直接讨论这个问题,我继续讨论 20 世纪的实验和相关的心理学理论。第一篇文章来自布卢门菲尔德(Blumenfeld,1913)。(参见表 1 的年表。)

布卢门菲尔德首先完成了一个特殊实验来表明,在某种意义上,现象学的视觉判断并不满足所有的欧几里得特性。布卢门菲尔德根据所谓的平行径(parallel alleys)和等距径(equidistance alley)做实验。在一间暗室里,受试者坐在桌旁,正视前方,然后,要求他调整放置在法面(即与连接两眼中心的水平线相垂直的平面)两边的两排光源。把两个最远的光源固

291

表 1 视觉空间是欧几里得空间吗?

人　　名	主　张	回　答
欧几里得(300 B. C)	透视理论	是
里德(1764),丹尼尔斯(1972),安格尔(1974)	视觉几何是球面几何	不是
布卢门菲尔德(1913)	平行径不等于等距径	不是

（续表）

人　　名	主　　张	回　答
吕内堡（Luneburg, 1947, 1948, 1950）	视觉空间是双曲空间	不是
布兰克（Blank, 1953, 1957, 1958a, b, 1961）	本质上与吕内堡一样	不是
哈代（Hardy, 1953）等	本质上与吕内堡一样	不是
扎永奇科夫斯卡（Zajaczkowska, 1956）	关于吕内堡理论的实验检验的肯定结果	不是
谢林（Schelling, 1956）	相对于已知的固定点是双曲的	不是
戈格尔（Gogel 1956a, b, 1963, 1964a, b, 1965）	语境几何的等距趋势证据	不是
福莱（1964, 1965, 1966, 1969, 1972, 1978）	视觉空间是非均匀的	不是，除了
因东（Indow ，1967, 1968, 1974a, b, 1975）	MDS 方法产生好的 欧几里得拟合	不确定
因东等（1962a, b, 1963）	接近于因东	不确定
西川（Nishikawa, 1967）	接近于因东	不确定
松岛（Matsushima）和野口（Noguchi, 1967）	接近于因东	不确定
格林鲍姆（Grünbaum, 1963）	质疑吕内堡理论	是
斯特劳森（Strawson, 1966）	现象几何是欧几里得几何	是

定在对称的位置，并离法面的距离相等。接着，要求受试者排列其他光源，使这些光源从固定光源向他延伸形成一个平行径。292 他的任务是排列光源，以便他在自己的视觉空间内感到这些光

源是笔直的和相互平行的。这是作出一条平行径的任务。第二项任务是作出一条距离径。在这种情况下，除两个固定光源外，关掉所有的其他光源，只留下一对光源，调整这对光源使其离固定光源的物理距离是相等的——前面讨论过这种等距判断。然后，关掉这对光源，打开另一对离他更近的光源，进行调节，以此类推。虽然物理位型不一致，但是，在欧几里得几何中，直线是平行的，当且仅当，它们彼此离任何一条共有垂线的距离相等。在布卢门菲尔德实验中观察到的差异被看成是视觉空间不是欧几里得空间的证据。当你离开受试者时，在平行径和等距径的判断中，这些直线分叉了，不过，平行径情况下的夹角往往大于等距径情况下的夹角。吕内堡认为，当人们离开受试者时，径的分叉支持了他的假设：视觉空间是双曲空间。

事实上，吕内堡在 20 世纪 40 年代后期的一些出版物中，是"视觉空间是双曲空间"这种观点至今的最强有力的支持者。他与他的合作者同心协力，详细地阐明了双眼视觉的一种数学理论，同时，提供了一系列实验调查来检验该理论的基本原理。在许多方面，吕内堡的文章(1947)仍然是对双眼视觉理论的最详细的数学处理。吕内堡没有进行广泛的讨论，只限于常曲率的黎曼几何，以保持刚体运动，即刚体的自由运动性。吕内堡在自然的双眼视觉的一个坐标系中，以一种相当令人满意的形式，发展了常曲率黎曼空间，尽管缺乏清晰的公理化处理。在一个相互垂直的感觉坐标系 α、β 和 γ(这些希腊字母不是吕内堡所用的字母)中，能够通过下列等式表示线元 ds：

$$ds^2 = \frac{d\alpha^2 + d\beta^2 + d\gamma^2}{\left[1 + \frac{1}{4}K(\alpha^2 + \beta^2 + \gamma^2)\right]^2} \tag{1}$$

这里

对于欧几里得空间，$K = 0$，
对于双曲空间，$K < 0$，
对于椭圆空间，$K > 0$。

最新近的处理方法参见因东(1979)的工作。

重要的是承认，吕内堡的进路严格地说是视觉空间的心理学进路。这条进路假定，物理学的欧几里得空间的变换产生了心理的欧几里得空间。在这种意义上，这不是视觉空间公理的基本定性的进路。吕内堡从来没有慎重地或明确地考察过导致证明视觉空间是常曲率黎曼空间的更一般的原始假设。

293

在这些一般的发展之后，他转向详细地论证这种观点：视觉空间的适当的常曲率空间是双曲空间[在方程(1)中，$K < 0$]。这里不可能详述吕内堡的论证，但他把他的论证建立在三种主要考虑的基础上，所有这些考虑在视觉理论中都受到了很大的关注：首先，在曲线好像是直线的地方，从双眼单视界得到的数据是物理曲线的数据(关于这些现象的数据可追溯到亥姆霍兹之前的时代)；第二，关于平行性判断的这种径现象前面已提到过；第三，解释下列视觉失真小屋的判断：在这个小屋里，画出几条适当的透视线，其结果好像是矩形或规则图形。在后一种情况下，吕内堡利用了埃姆斯(A. Ames)的一些未发表的经典的和引人入胜的证明。参见埃姆斯的相关文章(1946)。这个领域的困难之一是，在后来的文献中，总是不能令人满意地分析吕内堡提出的与三个典型问题相关的这种详细的数学论证和定量论证。更确切地说，通常是提出不同类型的新数据来表明，不同的现象不赞成吕内堡的假设：视觉空间是双曲空间。迄今为止，这两方面的文献都很多。本章的其余部分提供的大量参考文献根本没有涵盖全部范围。

吕内堡于1949年逝世，但他生前的许多学生和合作者继承

了他的工作,也提供了额外的实验支持和数学论证来赞成他的观点。我特别要提到布兰克的文章(1953, 1957, 1958a, 1958b, 1961)和哈代等人的文章 (1953),尽管这不是一个详尽的名单。另一个肯定的实验检验是由扎永奇科夫斯卡提供的(1956)。

谢林(1956)同意吕内堡的观点,但修改了一个重要之处:即负曲率的度量——也就是说,吕内堡赞成的双曲空间的度量——本质上是瞬时度量。在一个已知时刻,眼睛有一个确定的定点,依照谢林的观点,相对于这个定点,吕内堡的理论可能是近似正确的,但这个理论的应用很有限,因为眼睛通常不停地转动,这个定点也不断地发生变化。变化的基本事实在任何一个完全适当的理论中都必须加以考虑。

戈格尔(1956a,b, 1963, 1964a,b, 1965)曾研究过所谓的等距倾向,或者,在本章的语境中,我们可以称为贝克莱倾向。记得贝克莱坚持认为,离观察者的距离根本不是一个视觉概念,而是来自触觉。戈格尔在没有讨论贝克莱的精确分析的前提下,提供了一个重要的证据:当没有其他线索时,有一种强烈的倾向是,把这些对象离观察者的距离看成是相等的。戈格尔的这些认真谨慎的研究,不仅对于确立等距倾向是重要的,而且,一方面对于确立个体变化的敏感性,另一方面对于确立额外的视觉线索的敏感性,都是重要的。等距倾向作为一种核心效应一定是现存的,但任何一个详尽的视觉空间理论都要解释令人困惑的语境差异和个体差异的复杂性,似乎对我来说,戈格尔的实验关于这一点在本质上是明确的。在所参考的论文中,戈格尔关于视觉空间特征的问题并没有给出明确的回答。但我已经把他列入表 1,因为似乎对我来说,他的研究的作用是极力论证对在标准的层次结构中确定很高层次的视觉空间的几何产生的

294

怀疑。更准确地说,他的研究所支持的观点是:充分的几何在特征上显然是取决于语境的,因而完全偏离了经典的层次结构。

视觉空间几何的许多令人感兴趣的实验研究是由约翰·福莱完成的。在福莱(1964)的文章中,他用小的点光源的有穷位型做实验,来检验视觉空间的笛萨格特性(Desarguesian peoperty)。[①] (当然,根据许多其他公理对视觉空间是有效的假设来检验这种特性。)对于大多数观察者(但不是全部)来说,这些结果确证了笛萨格特性。在福莱(1966)的文章中,把感觉到的等距(perceived equidistance)作为一个视距(viewing distance)函数来研究。像福莱的大多数实验一样,这个实验也是在视平线的水平面上进行的。在离观察者 1.2 米、2.2 米、3.2 米和 4.2 米的距离处,确定感知到的等距的轨迹(locus)。像在福莱的其他实验中那样,刺激物是在漆黑中看到的很小的类点光源。让观察者抬起头不动,但允许他的眼睛自由转动。有五个光源,一个光源固定在法面上,其他四个可变光源相对于法面以 12 度的角和 24 度的角固定在法面两边。在朝向观察者的所有距离点,感知到的等距的轨迹成凹形。也许最重要的是,这个轨迹会随视距的变化而变化,这表明,视觉空间不仅仅依赖于视网膜刺激的空间分布。此外,这里还有对语境几何的一个

---

① 首先,在说明笛萨格(Desargues)的命题时,我们需要两个定义。两个三角形是**来自一个点的透视图**,当且仅当,两个三角形的顶点是一一对应的,使得通过三对相对应的顶点的三条线相交于同一点,即透视点。两个三角形是**来自一条线的透视图**,当且仅当,两个三角形的三条边是一一对应的,使得构成三角形的三条边的三对相对应的线相交于位于同一条线上的点。笛萨格特性或命题是:两个三角形,是来自一个点的透视图,也来自一条线的透视图。在射影几何中,平面可能要么是笛萨格平面,要么不是,在三维射影几何中,这种特性根据其他标准公理是可证明的。

直接的论证,并且结果与吕内堡的理论不一致。等距判断具有下列类型。除了固定光源之外,命令一位受试者摆放每个光源,光源离法面的距离与观察者离固定光源的距离相等。因此,好像应该对他来说,光源位于以他自己作为观察者为圆心的一个圆周上。重要的观点是,在这个实验中的十位受试者中没有一个人断定,等距轨迹位于前面作为吕内堡理论的支持论据提到的维-苗二氏双眼单视界或环(Vieth-Mueller horopter or circle)上。对视觉空间的基本几何同样重要的是这样的事实:观察者确定的轨迹相对于法面是不对称的。

下面是福莱的另一个漂亮实验(1972)。情境如图 2 所示。受试者作为观察者位于 O 点,已知点 A 固定在 O 点的深度轴的正前方,命令受试者对垂直(⊥)和全等(≈)作出如下判断:

295

1. 找到点 B,以使 $AB \perp OB$ & $AB \approx OB$。
2. 找到点 C,以使 $OC \perp OB$ & $OC \approx OB$。
3. 判断 OA & BC 的相对长度。

结果是,在四十八次实验的四十位受试者中有二十四个人断定 BC 明显比 OA 长。BC 比 OA 长的这个判断与任何一个常曲率空间的特性相矛盾,不管曲率是正的、负的还是零。

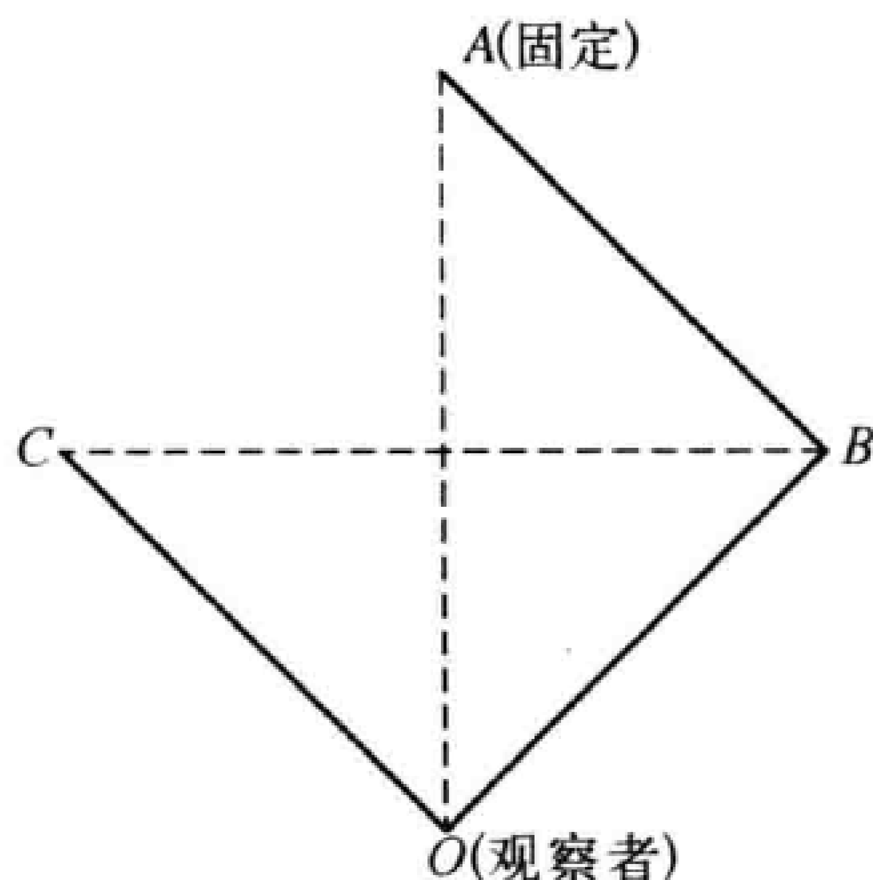


图 2 福莱的实验

福莱(1972)的实验研究表明,一方面,大小距离不变的假设是不正确的,事实上,正面的感知大小与以自我为中心的感知到的距离之比,远远大于物理比率,而另一方面,感知到的视角很接近于物理角度。这些结果和其他标

准假设都与吕内堡的视觉空间是双曲空间的理论不一致。福莱在这篇论文中以下列方式描述了第三个实验：

> 在最初视觉空间中感知的大小距离之比的报告，与以欧几里得的方式感知到的视角的报告无关，这如何可能呢？一种可能性是，两种判断在某种程度上是不同的和独立的知觉过程的产物……这些结果与两种判断是独立过程的产物这个假设相一致。它们也表明，没有一种几何模型能够适合于所有的刺激情境，而且，它们间接地表明，由于引入距离提示，这种几何可能接近于欧几里得几何。
>
> (Foley, 1972: p. 328)

此外，在福莱的分析中存在着对一种语境几何的强有力的辩护。这里一直没有提到的福莱的其他许多详细的实验研究，辩护了同样一般的语境观，下面我较详细地对此加以讨论。

因东太朗(Tarow Indow, 1967, 1968, 1974a, b, 1975)和与他有密切联系的其他日本研究者已经开展了有关视觉空间的几何的一系列详细研究(Indow et al., 1962a, 1962b, 1963; Matsushima and Noguchi, 1967; Nishikawa, 1967)。例如，他们发现，已经在心理学中得到了大力发展的多维标度法(MDS)，在许多情况下，产生了与欧几里得空间的极好的拟合(Indow, 1982)。因东根据实验检验基于可追溯到布卢门菲尔德(1913)的这种径实验的吕内堡的理论。如期所料，他重复了这种结果：等距径总是位于平行径之外，在其他的标准假设下，这个结果意味着这个空间的曲率是负的，因此，它必定是双曲空间。另一方面，因东(1974a, b)对吕内堡假设的简单性，尤其是曲率的恒定性假设，提出了真正的挑战。正是在这种语境中，他也尝试了确定多维标度法与欧几里得度量如何很好地拟合的替

296

代进路。正如他强调的,吕内堡的进路基本上以微分几何为基础,微分几何被看成是描述常曲率的黎曼空间的一种方法,但对于视觉判断来说,可能更加适当的是依靠大量判断,因而依靠视觉几何的不同概念基础。在他的整篇文章中,因东承认对视觉空间给出正确描述的任何一种简单答案的复杂性和困难。希望更深入地追究这些问题的读者,建议进一步详细地研读他的文章和他的合作者的文章。

我本人从因东及其同事的文章中比从其他人的文章中更多地了解了关于视觉空间的理论。我最初很怀疑吕内堡的思想,而且,正是因为因东及其同事完成了格外仔细的实验,我才意识到,有可能把吕内堡的这些思想转换成现实的方案。在这种情况下,这个方案并没有证明比原来更令人满意,这个事实不是因为这些实验无力,而事实上是因为这些实验太有力。它们使我们有信心承认,吕内堡的视觉空间进路在一些基本问题上是有误的。首先,因为因东及其同事在几种情况下已经指出,一旦我们假定,一个常曲率空间(例如,在双曲空间的情况下是负曲率)是一个合理的假设,就会完全缺乏参量的稳定性。当我们估算曲率时,我们发现,对于一个特定的受试者来说,即使是相同的实验,结果也是逐日变化的,无疑,当我们从一类实验(例如,经典的径实验)转到一个不同类型的实验的判断时,所估计的参量值很少能从一种情境转到下一种情境。

德勒斯勒(Jan Drösler)在一系列文章(1966,1979a,b,c,1987,1988,1992)中给出了对前面评论的视觉空间的几个不同特征的认真的数学分析。主题包括从凯莱-克莱因几何(Cayley-Klein geometries)中的多维标度法到双眼空间感知的心理物理函数(1988),在更丰富的心理学框架内,这个函数尤其与吕内堡的心理物理思想的更详细发展有关。

另一个实验(Wagner,1985)研究这样的感知判断：在户外的一个观察区的十三个白色木桩中,判断关于距离等量的测量。在物理坐标中,设 x 等于沿着深度轴的测量,y 等于沿着额轴的测量,再设距离的感知判断用上撇号表示(深度轴和额轴,参见图 3)。于是,瓦格纳的一般结果是,如果在物理上 $x = y$,那么 $x' \approx 0.5y'$。注意,严重的透视缩短是如何沿着深度轴发生的。瓦格纳的这个结果不是反常的或奇异的,而是代表了许多不同的感知实验：这些实验体现了沿深度轴的明显的透视缩短。 297

格林鲍姆(Grünbaum)在他关于时空哲学的重要著作中(1963),拒绝接受吕内堡的理论,并且证实,为了产生正确的感知判断,视觉空间必须是欧几里得空间。他的论证相当简洁,我将不作详细考察。我自己的观点是：他并没有在严格意义上给详细的实验研究足够的分量,或者,给已经提出的各种各样的理论建议的细节足够的分量。

我转向对这个问题的哲学回应来结束这种概述,即斯特劳森(Strawson,1966)在他关于康德的《纯粹理性批判》的书中的哲学回应。从我概述的大量的心理学文献的观点来看,令人惊奇的是发现斯特劳森把现象几何是欧几里得几何断言为必然命题。下面的引文坦率地陈述了这个问题：

> 根据后面要考虑的某些保留意见和条件,似乎也可以把欧几里得几何解释为关于现象的直线、三角形、圆等的一些不可证伪的命题;解释为关于这些类型的空间表象的先天命题,因而当然也解释为其应用限于这些表象的一种理论。
>
> (Strawson, 1966: p. 286)

斯特劳森的观点令人惊奇的特征是没有考虑这样的问题：现象几何可能不是欧几里得几何，而且，无论如何，确定具体情况是什么必定是一个经验研究问题。他后来给出的条件与这个问题无关，但与理想化的问题和作图的本性问题等非常相关。没打算以任何一种方式研究关于视觉空间本性的大量理论的和实验的文献资料，是令人吃惊的。

6.6 福莱和瓦格纳实验的部分公理[①]

我不能对这些实验中发现的得到强烈支持的实验事实给出完全令人满意的几何分析，但我确实认为，讨论一些有趣的问题有助于澄清基本情况。我把我建议的公理划分为各种不同的群。

仿射平面。我把这些看成定义 1.1 的标准公理。我们当然可以简化它们，也不要求是整个平面，但在这里那是不重要的。我们能够把介中性或平行性和像中点代数(midpoint algebra)之类的某个概念作为基元(primitives)。此外，这个决定不是这里考虑的关键。我们现在增加感知到的全等($\approx$)的判断。

三个特征点。直观上，设 o_1 是左眼的中心，o_2 是右眼的中心，o 是线段 o_1o_2 的平分点。显然，它们满足下面的两个公理。

298　2a. 这三个点是共线的和不同的。

2b. $o_1o \approx oo_2$。

全等的仿射公理。

3a. 任何一个平行四边形的对边全等。

① 这一节参考了苏佩斯文章(1995)中的资料。

3b. 如果 $aa \approx bc$，那么 $b = c$。

3c. $ab \approx ba$。

3d. 如果 $ab \approx cd$ & $ab \approx ef$，那么 $cd \approx ef$。

3e. 如果 $a|b|c$ & $a'|b'|c'$，$ab \parallel a'b'$ & $ab \approx a'b'$ & $bc \approx b'c'$，那么 $ac \approx a'c'$（这是一个熟悉的弱仿射加法公理）。

额轴是包括 o_1，o 和 o_2 的直线，**深度轴**是经过 o 的半直线，使得在这个轴上的任何一个点 a，$o_1a \approx o_2a$。（注意，我们不能根据一般的垂直概念描述深度轴，因为垂直概念没有用，并且事实上，在这些公理框架内，也将没有用。深度轴只是一条半直线，因为一个受试者不可能直接看到额轴后面。）

3f. **第一个特殊的全等公理**。如果 $a \neq c$，a 和 c 在额轴上，b 在深度轴上，并且 $ao \approx oc$，那么 $ab \approx bc$（见图 3）。

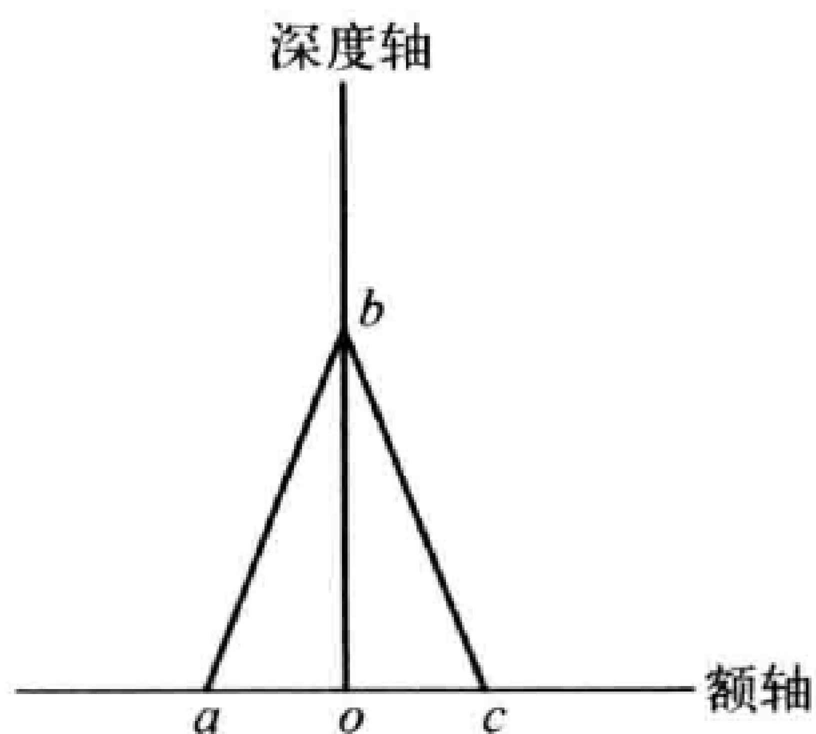


图 3 全等公理 3f

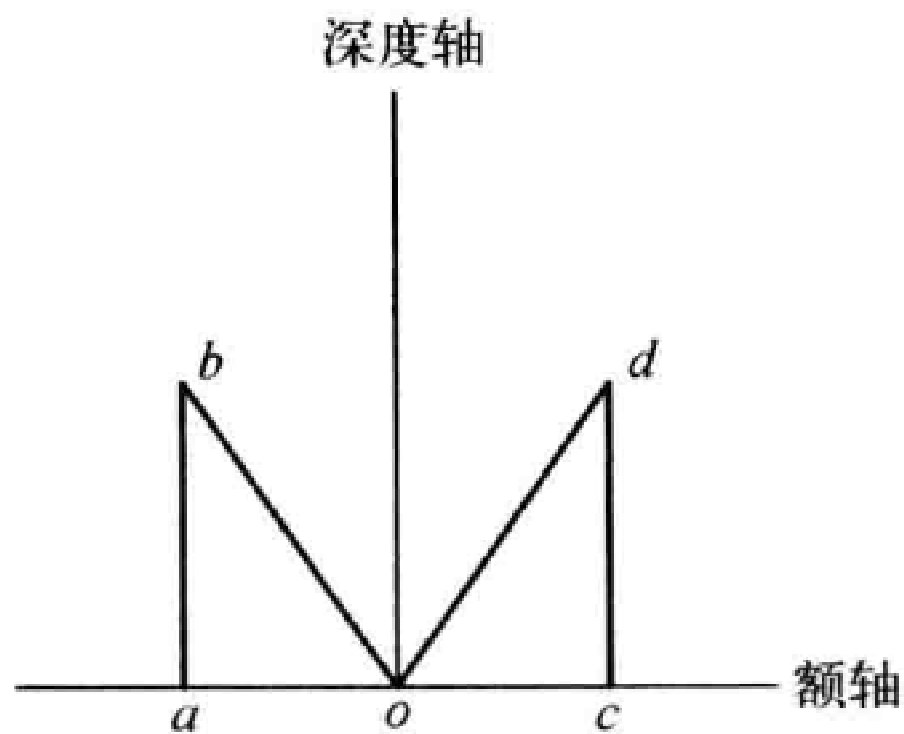


图 4 全等公理 3g

3g. **第二个特殊的全等公理**。如果 $a \neq c$，a，c 在额轴上，$ao \approx oc$，ab 和 cd 平行于深度轴，$ab \approx cd$，那么 $ob \approx od$（见图 4）。

后面两个特殊的全等公理把仿射全等推广到不平行的线段的全等，而且在这种情况下，如图 3 和图 4 所示，线段相对于深度轴是对称的。这意味着，我们有一种仿射全等的弱扩展，一种

扩展太弱,就不能为我们提供绝对空间的全等公理(参见苏佩斯等,1989,第13章)。(正如已经评论的那样,能够这样考虑绝对几何。放弃下列欧几里得公理:通过一条线 α 之外的一个已知
299 点 a,至多有一条线平行于 α,并位于由 a 和 α 形成的平面内。把这个公理加到绝对几何的公理中,为我们提供了欧几里得几何,显然是这样。更有趣得多的是,把这个公理的否定加到绝对几何的公理中,为我们提供了双曲几何。)我们能证明下列定理。

定理1. 当深度轴是指前半平面时,设半空间是由额轴的同侧的所有点组成的:

(1) 前半平面,在实值函数 φ 下,同构于实数域上的二维仿射半平面,其中,x 轴为深度轴,y 轴为额轴。此外,当线段全等时,作为通过欧几里得度量测出的全等得以满足,即,如果 $ab \approx cd$,那么 $\sum_{i=1}^{2}(\varphi_i(a)-\varphi_i(b))^2=\sum_{i=1}^{2}(\varphi_i(c)-\varphi_i(d))^2$。

(2) 惟一可能的仿射变换是由额轴的伸长 α 和深度轴的伸长 β 构成的那些变换,其中 $\alpha,\ \beta>0$。

从经典几何的熟悉结果来看,(1)的证明是显而易见的。(2)的证明可从下列观察推出:所描述的仿射变换是围绕两个特殊的全等公理要求的深度轴保持对称的仅有的变换。

显然,定理(1)的结果是相当弱的。福莱和瓦格纳的实验并没有否定这个定理,但这是不足为奇的,因为介中性方法加上关于深度轴的对称全等不可能描述任何一个实验的结果。正如福莱的步骤所要求的,如果我们增加一个垂直概念,那么我们本质上会得到一个欧几里得半平面,由此而得到的结构将否定福莱的结果。

相应地,如果我们不增加与吕内堡的那些假设完全不同的一些心理物理假设,就不能描述瓦格纳的沿深度轴的广延的感

知的透视缩短的心理物理结果。当然，福莱的结果也最好被解释为是沿着深度轴的感知的透视缩短。自然的结论是，在接近于常曲率空间结构的特殊性和简单性的任何空间内，我们不可能一致地描述视觉几何关系。 300

但根据上面给出的弱化的全等公理，甚至定理 1 的弱仿射结构也太强，大概应该用一个标准版本的绝对几何取而代之。这样一种绝对几何能被扩展到双曲几何，但仿射结构则不能。我们似乎以描述不同实验的各种不同几何结构的片段而告终。双曲几何片段也许适合径实验，标准几何之外的片段适合福莱的实验，等等。

现在，企图用一种统一的视觉空间结构来充分说明所有重要的实验结果的目标，似乎是错误的。只有一种多元的和片段的进路似乎才是可能的。

6.7 关于视觉空间的三个概念问题

这一节我集中评论三类问题。首先关注视觉几何的语境特征，其次是关于距离感知和运动的问题，第三是关注描述视觉空间的对象本性的问题。

语境几何。大量不同的实验和日常经验证明了视觉空间的高度的语境特征。有没有“额外的”点可能会极大地影响到感知判断。整个错视范围(我至此还没有讨论这个问题)也为这些语境效应的惊人的说服力提供了大量证据。

就我所知，还没有人试图站在视觉几何的公理基础的立场上认真地考虑过这些语境效应。在某种程度上这是正常的，因为公理化基础隐含的意义，从通常的观点来看，是可怕的。让我

们举一个简单的例子来说明这一点。

在普通的欧几里得几何中,只有当三角形的两条边有相同的长度时,三个点才能形成一个等腰三角形。现在假定,欧几里得几何有更加复杂的方面:一个三角形是否是等腰三角形,不仅依赖于这三个点的位型,而且依赖于是否有一个特征点恰好位于这个三角形的一条等腰边的外侧。这种不对称完全能使视觉三角形不再看起来是等腰三角形。这只是语境效应的可怕组合的一个简单事例,这种语境效应,很容易能被想象出来,而且,不需要丰富的想象力或实验技能,就能被证实为是真的。关于这个主题的历史和令人困惑的各种不同的错视事例可参见科伦(Coren)和吉尔古斯(Girgus)的合著(1978)。

关于这些效应我们能说些什么呢?似乎对我而言,最重要的问题是承认,感知的几何实际上完全不同于经典几何,但根据我们作出的这些判断,它更接近于物理学。例如,考虑两个物体通过万有引力彼此吸引的相应情境。引入第三个物体使原来两体的运动变得完全不同,对于别的情形来说,这被认为是古怪的。这也适用于作用的电磁力、机械力,等等。语境效应在物理学中是常见的,并且,物理学家建立相关的物理学理论说明这些效应。

301

注意,物理学理论依赖于位于时空中的特殊位置的特征对象。时空本身是非特征点的一个连续统,下面这点是经典几何的公理化基础的典型特征:即在这个空间中不存在非特征点。但我们总是研究类似于物体的特征点,而不是几何点,这恰好是一个感知的特征。鉴于这种观点,我们说我们在视觉几何中有语境效应,就像我们把广义相对论中作出的类似主张归因于已知区域中有大质量物体一样自然。

足够有趣的是,有证据表明,当我们增加视觉线索时,也就

是说，当我们用越来越复杂的视觉想象语境填补视觉域时，视觉空间越来越成为欧几里得空间。我们这里的情境有可能与广义相对论中存在的情境正相反。在感知的情况下可能是，很容易使由少数几个视点组成的空间不同于任何一种标准几何。

当把在某一标准几何中的有穷点集的同构嵌入当作是对视觉空间本性的适当分析时，就能使几何学的观点与物理学的观点密切地结合起来。这条进路在前面提到过，并隐含在所讨论的一些实验文献中。它没有被充分地揭示出来，并且，在一个已知空间里，不管是欧几里得空间、双曲空间、椭圆空间，还是无论哪一种空间，一个有穷点集的嵌入，必须以惟一的方式满足的那些定性公理的整个范围，都需要引起更明确的和更详细的关注。

似乎同样令人满意的是，在最初的这类研究中，通过谨慎地引入对称性还有某些额外假设，比如，固定的类型与观察者之间的很特殊的关系，来避免语境效应问题。许多不同的实验研究和源于吕内堡传统的那种数学分析表明，在特殊限定的假设前提下，能够达到许多肯定的和几乎确定的结果。同时，使这些结果尽可能成为明确的，同时承认，这些结果的独特特征和接受一般情境在特征上是语境的这个事实，似乎是一项适当的研究策略。针对这些特殊情境，人们似乎也有可能对“视觉空间是欧几里得空间吗?”这个问题给出一种断然否定的回答，而且，以很高的近似度回应说，在许多特殊情况下，视觉空间是双曲空间，在其他特定的情况下，视觉空间在特征上是椭圆空间。这种受限制的答案肯定是否定的。至于如何以一种完全满意的方式描述这种几何：即它说明了作为错视、等距倾向等的典型特征的语境效应，目前的一般回答似乎是无用的。

距离感知和运动。正如前面在简要讨论亥姆霍兹—李问题时所表明的，上一节概述的大多数工作没有充分地说明运动问 302

题。福莱有一篇关于距离感知的极好的概述性文章(1978),这篇文章表明,在最初聚焦于一个对象期间,在获得关于感知距离的信息时,眼动是特别关键的。感知问题的哲学传统往往忽略了作为视知觉的一个整体部分的眼动或头动的复杂问题,但最初步的考虑就足以表明它们的基本重要性。吕内堡的一个基本的远见,是承认描述互补的眼动和头动的不变特性是重要的。对确定视域特征的更深层的跳读问题,实际上,还没有以一种完全数学的和定量的方式研究过。依我的看法,这无疑是视觉空间理论未来发展的最重要的领域。我设想,我们应该用视知觉的运动学取代视知觉的几何。例如,兰姆(Lamb)证明(1919),东德斯(Donders)定律断言,眼球的位置完全是由初始位置和对准固定点的视轴决定的,根据这个定律,不可能把每一条物理线段都看成是直的。在罗伯茨(Roberts)和苏佩斯的文章(1967)中曾详细地阐述过兰姆的这个运动学定理,这个定理提供了反对视觉空间的欧几里得特征的一个有力的运动学论证。我这里引证它,只是作为人们在更彻底地发展视知觉的运动学时应该期待获得的这类结果的一个例子。

视觉空间的对象。本章给出的整个分析,并没有以任何一种精确的或明确的方式澄清视觉空间对象的准确描述是什么这个问题。这种含混性是故意的,因为我所参考的广泛文献关于什么将被视为视觉对象并没有一种不变的说明。包括的观点是极端不同的——从甚至不愿承认纯视觉空间几何的贝克莱,到坚持认为,视觉空间只是一种标准的欧几里得空间和在视觉对象和物理对象之间根本没有真正区别的那些人。在公理化和系统化地建设这门学科时,显然需要有些约定,然而似乎是人们能够在不固定于一种精确描述的前提下,对所考虑的文献范围进行了清楚的讨论,因为在视域中看到的东西基本上是一致的。

站在视觉空间几何的立场上，为了讨论几何的特征，我们甚至能允许关于对象是二维的还是三维的如此宽泛的争论。托马斯·里德强烈地倾向于视觉空间的二维特征。福莱坚持认为，视觉空间是三维的；然而注意，他的大多数实验限于二维。最起码，根据对视觉空间对象的几种不同的自然描述，显然能够作出的强主张是：视觉空间不是欧几里得空间，而且，这是具有哲学趣味的一个结论。

6.8 几何学中的有穷论[①]

在20世纪数学基础的发展中一直占有统治地位的是，关注为经典分析和经典数论提供适当的概念基础。弗雷格、希尔伯特、罗素(Russell)和布劳维尔(Brouwer)及他们的追随者的态度，一直集中于为纯数学的经典部分提供严格的概念基础。关于这段历史值得关注的是，对应用数学中的有穷论是多么不重视。对各种不同应用传统的关注将支持比过去可能的情形更基本的有穷论，只要这种观点是为经典的初等数论提供一个适当基础的观点。即使最弱的递归算法系统也有闭合条件，这些闭合条件使得这些作图远比任何标准应用中的作图要求复杂得多。对闭合性和完备性的热爱本身是极好的智力热情，但是，这些热情与通过在计算上可行的极限方法求解应用题的动力背道而驰。当把数学应用于真实的世界时，这条思路很容易导致有穷论。 303

那些关注纯数学的人通过说在纯几何中对有穷论已经有大量的强调，然后，通过引证近半个世纪以来关于有穷几何的广泛

---

① 这一节的资料取自苏佩斯的文章(2001a)。

而深入的研究,可能立即回应我对有穷论的强调。把这项工作看成是对费利克斯·克莱因的埃尔兰根纲领(Erlangen program)——即首先强调几何的自同构群结构的几何进路①——的一种推广,也是自然的。自同构的研究也是研究有穷几何的一个重要方面,因为已知一种有穷几何,自同构群是一个有穷群。同样已知一个有穷群,自然的问题是确定它的有穷几何。正如前面所注意的,直到20世纪60年代中期,登博夫斯基(Dembowski, 1968)才给出一种综述。近三十年来也有大量的工作,而且,解决了纯数学类的许多有争议的问题,尤其是有穷几何与有穷单群的关系方面的问题。

然而,这些大量的研究不是本节的主题。我想到的是聚焦于更基本的应用数学问题。我想到的应用是什么呢?我以自古以来建筑师所用的数学方法为出发点。早在公元前6世纪,建筑师就把几何应用于详细的作图问题。他们并不是简单机械地应用从当时的哲学和数学传统来看很熟悉的初等几何,而是他们运用详细的几何思想来提供视错觉。提供视错觉的这些方法导致了在当时和现在的建筑学文献中所谓的"精致"。一个重要例子就是**凸肚状**(entasis),这是一个使大理石柱垂直的轮廓凸出来而不是直线的几何作图。维特鲁威斯(Vitruvius)在公元前1世纪关于这种作图是这么说的。

304

> 由于高度不同,随着眼睛视线的上升,在柱的粗处,产生了有比例的放大。因为视觉总是追求美观,并且,如果我们不能通过在尺寸上的有比例的放大来满足达到快乐的欲望,因此弥补视觉错误,就会给观众呈现出不雅而难看的外

---

① 参见4.3节。

观。在希腊人中间，把柱子中部的放大称为恩塔西斯（εντασις），书末附加了一幅图和计算结果，表明了如何据此产生出优雅和适宜的效果。

（Vitruvius，1960 edition：p. 86）

当然，凸肚状是一个著名的例子，但还有更多的事例；现存最长的书面记载是维特鲁威斯的罗马文著作，写于奥古斯都时代。维特鲁威斯书中描述的详细应用并不奇异，但标准。令人有理由遗憾的是，帕台农神庙的建筑师伊克蒂诺（Ictinus）所写的建筑学的书现在遗失了。在当前语境中重要的一点是，在伊克蒂诺很久之前和维特鲁威斯很久之后，建筑师要接受在几何和如何细致地把几何应用于建筑施工图的训练。正如我喜欢指出的，这些建筑师不是在证明定理，而是在画图（欧几里得语言中的问题）。这些作图并非全是简单的和显而易见的；事实上，丢失的维特鲁威斯对凸肚状图的画法，在其数学内容上，比今天在最著名的建筑学著作中所能找到的画法更复杂。（但肯定没有现代建筑设计的计算机程序的数学背景那么复杂。）正如在帕拉迪奥的许多阐述中可以看到的，维特鲁威斯之后的一千多年，在帕拉迪奥的四本著名的建筑学书中（1570/1965），希腊的比例传统，而不是直接的数字计算，一直统治着建筑学的详细思维。现在可用的计算机资源，以各种不同的方式，再一次鼓励强调比例。

除了凸肚状和其他视错觉之外，公元前5世纪的建筑师所用的最重要的数学或几何发现是透视法的发现，传统上归功于画家阿加萨霍斯（Agatharchus），他曾为埃斯库罗斯（Aeschylus）的悲剧设计舞台背景，结果还以此为经验写了一本关于布景制作的书。透视理论的数学研究最早开始于德谟克利特和阿那克萨戈拉。这就是在五百年后维特鲁威斯写于罗马的著作中所说

的关于透视的早期著作。①

> 在雅典，当埃斯库罗斯出版了一本悲剧体裁的作品时，阿加萨霍斯曾为此布置场景并留下了解说词。这致使德谟克利特和阿那克萨戈拉撰写同一主题来表明，已知一个确定地点的中心，直线应该如何与自然注意到的视点和视线的偏离相呼应，以便根据这种错觉，在布景中给出建筑物外观的可靠表征，以及以便画在一个垂直平面上的正面建筑，有些部分看上去缩进背景，而另一些部分看上去凸了出来。
>
> (Vitruvius, 1960 edition: p. 198)

305 关于几何的这些不同应用(当然包括重要的透视情况)所重要的是，方法和结果在特征上是高度有穷的，事实上，恰好像希腊几何的大多数(如果不是全部的话)结果一样。

因此，如果我们在包括透视数学理论和射影几何的整个发展在内的这种悠久而重要的建筑学传统中，看待几何的应用，人们自然对数学的有穷论特征和熟悉的应用之间密切的相互影响感到震惊。古代数值方法的应用有相同的有穷论特征，但我在这里不打算考虑细节。

在19世纪的一个特定阶段，由于几种不同的力量在起作用，非有穷论方法用于经典分析并在某种意义上被推广到几何。我特别强调需要有完备形式的表征定理，因为开始阐述几何空间的表征，是为了同构于已知的实数系统的笛卡儿积，这个积本身取决于空间的维度。对表征定理的这种研究根本不是更早时期的希腊传统的一部分。根据对数学基础的许多不同讨论，提

---

① 根据亚历山大·琼斯(Alexander Jones)的观点(2000)，在古代，只有两位关于数学光学的作者研究过透视问题，一位是托勒密，在他的著作《地理学》第七卷中，另一位是帕普斯(Pappus)，在他的著作《选集》第六卷中。

出这个问题的一种方式是,这样的表征定理要求一种实无穷的表征。我们正是把这种实无穷的表征,或用更书面的术语来说,是无穷基数集的表征,从几何空间映射到在同构意义上保留了几何空间结构的实数集的一个笛卡儿积。包括选择定理或其等价方法在内的非有穷方法的更强的意义是,更进一步远离了我这里所关注的应用的有穷论。

一个不同的问题是,在物理学等学科中应用数学时,在多大程度上运用经典分析,才能不仅避免承诺运用选择公理的任何结果,而且避免承诺一个实无穷特征的任何集合。① 我自己深信,人们能够以纯有穷论的方式坚持到底,或肯定能几乎坚持到底,但这不是我这里要关注的。

无量词公理与作图。很严格的作图方法的一个必要条件是,所讨论的理论的公理应该是无量词的,因而避免所有的纯存在假定(purely existential assumptions)。这样的存在要求被特殊的作图所取代,根据已知的原始作图,这些作图要么是原始的,要么是以无量词的方式可定义的。通常,在传统的几何表述中,存在公理在下列意义上是无害的:它们能够很容易被基本上等价的无量词公理所取代,这些无量词公理运用了固定的几个已知常数。

时序。在作图的标准公理化表述中所忽略的一个应用问题是完成任何一个实际作图所需的时序。通常运算或关系的数学 306

① 19 世纪使分析算术化的基本努力甚至迫使最基本的概念也有一个作为外延的无穷集,例如,把一个已知有理数表征为所有分数的等价类,在通常意义上,相当于一个比例分数(representative fraction)。因此,有理数 $\left[\frac{1}{2}\right]=\left\{\frac{p}{q}: 2p = q \ \& \ q \neq 0\right\}$。但这种非直观的抽象在应用中从来没有起到过重要作用,也没有在证明标准的分析定理时发挥过作用。

符号没有地方表明时间顺序。然而,在概率论中,这样的时间问题,在发展随机过程的理论或应用中,具有很大的重要性。不管把时间看成是连续的,还是离散的,事件或随机变量可能发生的时间用下角标标出,用 t, t', t_1 等表示连续时间,用 n, n', $n+1$ 等表示离散时间。当可能发生的时间没有被规定为是理论框架的一部分时,那么,原则上,需要一个新的随机变量来描述这种不确定性。

这样的概率考虑不在这里考虑的几何框架之内,虽然如果在实际作图时考虑到实际的误差事实,自然能把它们包括在内。这样的问题在应用中无疑具有重要性,但这里只是为了简易起见而忽略这些问题。这是因为从一开始就包括了在实际作图时处理不可避免的近似误差的概念方法的公理的更大复杂性。关于这些公理的复杂性的例子,参见苏佩斯等人的著作(1989,第16章)。

我们在作图符号中能容易用下角标表明时序。这提供了一种容易的方式证明,两种作图法是否需要相同的步骤。下面是来自仿射几何的一个简单例子:只用两种作图法找中点,即平分一条线段和把一条线段沿任一方向延长一倍。① 在直观讨论中,以习惯的方式提到线段,但在形式上要求只有两个点,即一条线段的两个终点。在这个例子中,我非形式地表述作图的公理。

问题。**已知三个非共线的点 α_0, β_0 和 γ_0,作一个邻边是 $\alpha_0\beta_0$ 和 $\alpha_0\gamma_0$ 的平行四边形。**

1. 作 $\beta_0\gamma_0$ 的中点 a_1。

2. 沿着 a_1 的方向把线段 $\alpha_0 a_1$ 延长一倍作点 a_2。图

① 即后面所说的"加倍"。——译者

$\alpha_0\gamma_0 a_2\beta_0$就是所要求的平行四边形(见图5)。

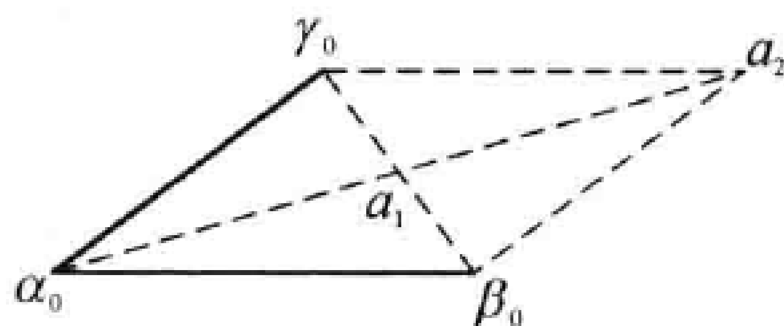


图5　作平行四边形

仿射公理。我现在转向仿射平面作图的公理,它们是通过平分或加倍点对之间的距离进行的。正如从我前面的评论中所预期的那样,这些公理是无量词的。在陈述这些公理时,我没有用到作图的时间顺序符号,但在作图时,如果需要的话,能够以明显的方式增加适当的下角标。在陈述公理时,这些下角标不太有用。 307

一些基本定义和定理采用了这些公理,但是,除了引入坐标的表征定理之外,这里没有给出证明。(苏佩斯的文章中有更多的定理和证明,虽然所用的符号稍微有些不同;Suppes, 2000。)为了避免关于运算的闭合条件,平分作图在形式上用关系$B(ab, c)$表示,其直观意思是,c是平分线段ab作出的点。相应地,$D(ab, c)$是沿着b方向把线段延长一倍的作图法,直观上,c的位置如图6所示。还需要作为基元的三元关系是线性关系。在直观上,如果a、b、c在一条直线上,那么$L(abc)$。

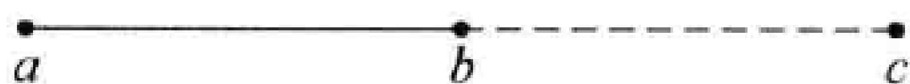


图6　加倍的作图法

首先,只含有线性关系的公理。

L1. 如果$a=b$, $a=c$或$b=c$,那么$L(abc)$。

L2. 如果$a\neq b$, $L(abp)$, $L(abq)$且$L(abr)$,那么$L(pqr)$。

其次,只含有平分的公理。注意,第一个公理断言,当点a和b(不一定是不同的)是已知的时,那么所作出的点是惟一的。

B1. 如果$B(ab, c)$和$B(ab, c)'$,那么$c=c'$。(惟一性)

B2. $B(aa, a)$。(幂等性)

B3. 如果 $B(ab, c)$,那么 $B(ba, c)$。(交换性)

B4. 如果 $B(ab, c)$, $B(de, f)$, $B(ad, g)$, $B(be, h)$, $B(cf, i)$和 $B(gh, j)$,那么 $i = j$。(双向交换性)

B5. 如果 $B(ab, c)$和 $B(ab', c)$,那么 $b = b'$。(消去性)

公理 B4 很好地举例说明了运用代数运算符号所带来的理解的简单化。于是,这个公理有更简单的形式:

$$(a \oplus b) \oplus (d \oplus e) = (a \oplus d) \oplus (b \oplus e)$$

正如从我的讨论应该明确的那样,我用笨拙的关系式符号使得对有穷位形的限制后来变得更加一目了然。

四个加倍作图的公理如下:

D1. 如果 $D(ab, c)$且 $D(ab, c')$,那么 $c = c'$。(惟一性)

D2. 如果 $D(ab, c)$且 $D(ba, c)$,那么 $a = b$。(反对称性)

D3. 如果 $D(ab, c)$且 $D(ab', c)$,那么 $b = b'$。(左消去)

D4. 如果 $D(ab, c)$且 $D(a'b, c)$,那么 $a = a'$。(右消去)

最后,有三个公理用到了不止一种关系。

BD. 如果 $D(ab, c)$,那么 $B(ac, b)$。(归纳)

308 LB. 如果 $B(ab, c)$,那么$L(abc)$。(线性平分)

LBL. 如果 $B(ab, d)$,$B(bc, e)$,$B(ac, f)$和 $L(def)$,那么 $L(abc)$。(线性中点)

我引入两个条件定义,来重新提出简化的运算符号。

定义 1. 如果 $B(ab, c)$,那么 $a \oplus b = c$。

定义 2. 如果 $D(ab, c)$,那么 $a * b = c$。

换言之,$\oplus$是平分的条件运算,$*$是加倍的条件运算。在下面运用运算符号的定理和定义中,假定利用$\oplus$和$*$运算的定义 1 和 2 要求的条件,在所有情况下都被满足,但没有明确地

陈述。

定理。首先,我在一个定理中总结了共线的基本特性。

定理 1.(共线)

(i) 如果 $L(abc)$,那么,对于 abc 的任意排列,L 都成立。

(ii) $L(aba)$。

(iii) 如果 $a \neq b$, $L(abc)$和 $L(abd)$,那么 $L(acd)$。

(iv) 如果 $p \neq q$, $L(abp)$, $L(abq)$和 $L(pqr)$,那么 $L(abr)$。

斯米卢(Szmielew, 1983)指出,定理的(i)、(ii)和(iii)等价于公理 L2 和 L3——实际上是(i)的一种更弱的形式,即,如果 $L(abc)$,那么 $L(bac)$。

定理 2. $(a \oplus b) \oplus c = (a \oplus c) \oplus (b \oplus c)$(自分配性)

定理 3. 如果 $a \oplus b = a$,那么 $a = b$。

下面的定理表明,对于加倍,归纳和线性成立。

定理 4. $a * (a \oplus b) = b$ 和 $L(ab(a * b))$。

任意三个不共线的点"形成"一个三角形,因此,我们可以把成为三角形(triangularity)的三元关系 T 定义为 L 的否定。

定义 1. $T(abc)$,当且仅当,它不是 $L(abc)$的情形。

在下一个定义中,所定义的四元关系 P 具有的直观意思是:在这种关系中排列的四个点形成一个平行四边形(因此,"P"表示"平行四边形")。

定义 2. $P(abcd)$,当且仅当,$T(abc)$和 $a \oplus c = b \oplus d$。

这个定义描述了平行四边形的特征,在标准的无法作图的(nonconstructive)仿射几何中,它"局部地"起到了平行线的作用。成为三角形或非线性条件排除了简并。重要的条件是,一个凸面四边形 $abcd$(见图 7)是一个

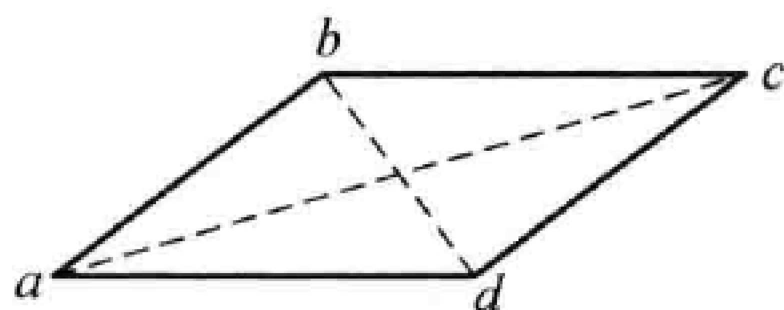


图 7　一个平行四边形的对角线的交点

平行四边形,当且仅当,对角线 ac 和 bd 的中点恰好重合,这是一个熟悉的平行四边形的特性,而且对这个定义来说,是充分的。凹面四边形违背中点条件。

我现在根据已经引入的概念非形式地定义什么是作图。已知三个非共线点 α_0、β_0 和 γ_0,作图是项(C_1, …, C_n)的一个有穷序列,其中,每一项 C_k($k = 1$, …, n) 都是一个有序的三元组:

309 (i) 倘若 a_k 既不是 α_0、β_0、γ_0,也不是前面作出的一个点,它的第一项是一个点 a_k;①

(ii) 它的第二项是一个点对(a_i, a_j),它要么是 α_0、β_0、γ_0,要么是前面按顺序(即 i, $j < k$) 作出的点;

(iii) 它的第三项要么是代表平分法的"B",要么是代表加倍法的"D",使得把所示的运算应用于点对(a_i, a_j)作出 a_k。

因此,前面以 $\alpha_0\beta_0$ 和 $\alpha_0\gamma_0$ 为邻边的平行四边形的作图是序列

$$((a_1, \beta_0\gamma_0, B), (a_2, \alpha_0 a_1, D))$$

正如已经评论的那样,作图要求在每一步都作出一个新的点。证实一个点不是已有点的作图方法是,在每一步都归纳地引入坐标,使得简单的数值检验成为可能。在许多应用中,这种理论上的坐标检查通常被每一步作图的快速目测所取代。

解析表征定理。下面是有穷表征定理。

定理 5. 已知三个非共线的点 α_0、β_0、γ_0 和基于这三个点的一个作图 $C = (C_1, \cdots, C_n)$,那么作为下列定义的有理数值函

---

① a_k 不是前面已知的或作出的,这个条件实际上是不需要的。这样一种条件,例如,在证明的通常的形式递归定义中是不标准的。另一方面,它是一种直观上的自然约束。

数对(φ_1，φ_2)提供了通过 C 作出的点(a_1，…，a_k)的解析表征，这里：

(i) $\varphi_1(\alpha_0)=0$，$\varphi_2(\alpha_0)=0$，$\varphi_1(\beta_0)=1$，$\varphi_2(\beta_0)=0$，$\varphi_1(\gamma_0)=0$，$\varphi_2(\gamma_0)=1$；

(ii) $\varphi_i(a \oplus b)=\dfrac{\varphi_i(a)+\varphi_i(b)}{2}$，$i=1, 2$（平分）；

(iii) $\varphi_i(a * b)=2\varphi_i(b)-\varphi_i(a)$，$i=1, 2$（加倍）。

大概证明：根据归纳法，对于 $n=1$ 来说，如果 $C_1=(a_1, \alpha_0\beta_0, B)$，那么

$$\varphi_1(a_1)=\frac{0+1}{2}=\frac{1}{2}$$

$$\varphi_2(a_1)=\frac{0+0}{2}=0$$

C_1 其余的 11 种可能情形是类似的。

如果对于 1，…，$n-1$，这个表征成立，那么，虽然对于已知的作图来说有许多可能的 C_n，但 C_n 是惟一的，由于用在归纳意义上定义的直到 $n-1$ 的坐标来检查，作出的点 a_n 是新的。因此，如果一般项具有形式 $C_n=(a_n, a_ia_j, B)$ $(i, j<n)$，那么 310

$$\varphi_1(a_n)=\frac{\varphi_1(a_i)+\varphi_1(a_j)}{2}$$

$$\varphi_2(a_n)=\frac{\varphi_2(a_i)+\varphi_2(a_j)}{2}$$

如果一般形式是 $C_n=(a_n, a_ia_j, D)$ $(i, j<n)$，那么

$$\varphi_1(a_n)=2\varphi_1(a_j)-\varphi_1(a_i)$$

$$\varphi_2(a_n)=2\varphi_2(a_j)-\varphi_2(a_i)$$

显然,赋予数值坐标的归纳法是显而易见的,证明也是如此,这是标准的几何作图应有的方式,与整个实平面或三维空间的仿射表征、射影表征或欧几里得表征的冗长而复杂的必要证明相反。在博尔舒克和斯米卢的著作(Borsuk and Szmielew, 1960)中,这样一种证明的一个严格的例子长达好几页。

解析不变性定理。还有一个有穷仿射不变性定理,对坐标的引入成立,我只陈述这个定理,没有证明。

定理 6. (不变性定理) 设 C 是基于三个非共线的点 α_0、β_0、γ_0 的一种作图,再设(φ_1, φ_2)是如上定义的坐标函数。
设

$$\varphi'_1 = a\varphi_1 + b,\ a \neq 0$$

$$\varphi'_2 = c\varphi_2 + d,\ c \neq 0$$

那么(φ_1', φ_2')也是 C 的一个坐标表征,C 随原点的可能变化而变化(如果 $b \neq 0$ 或 $d \neq 0$),也随仿射比例的变化而变化,除非 $a = 1$ 或 $c = 1$,现在,在这里,$\varphi'_1(\alpha_0) = b$、$\varphi'_2(\alpha_0) = d$、$\varphi_1(\beta_0) = a + b$、$\varphi_2(B_0) = d$、$\varphi_1(\gamma_0) = b$ 和 $\varphi_2(\gamma_0) = c + d$。

应该显而易见的是,对于每个特殊的作图来说,还需要更多的东西,即这样一种证明,在欧几里得几何中这种证明是如此熟悉,以至于这种作图实际上是在落实作图的原意。作图如同证明一样,每一步都可能是有效的,但是,如何作图,像如何证明一样,并非是为了作为问题或定理来给出的。因此,每一个作图都伴随着一个相匹配的证明,可能是直接的,也可能是运用解析表征。

第二,一般情况下,我们希望为路径无关的定理作图。如果我们给出点的一个有穷位型,那么,我们能证明,这个位型具有独立于用来产生它的特殊作图的某些特性。

特殊作图的这些附加证明的必要条件都不可能使我们超出已经描述的有穷论框架的范围。在苏佩斯的文章(Suppes, 2000)中,引入一个附加的梯形作图法,允许任何有理数的容易的有穷论的作图。

代数问题。强调通过作图引入坐标,极大地简化了解析表征定理。一条纯代数进路,含有对这些公理的有穷模型的一般表征,但没有给出点的作图顺序,这条进路要求把代数恒等式作为公理和更复杂的表征定理。所需要的这类等式的例子如下,这个例子是不完备的: 311

$$(a*b)*b=a$$

$$(a*b)\oplus(b*a)=a\oplus b$$

$$(a*b)*(a\oplus b)=b*a$$

很容易检查,预期的坐标表征要求满足这些等式。同样明显的是,这些等式和类似的等式代表了多余的作图步骤,根据作出的每个点都是新的点这个必要条件,排除了这些多余的步骤。这些等式足以很好地表明,为什么作图的有限性极大地简化了这个理论,这应该是打算在不同应用中所用的任何一个理论的目标,正如历史上领先于实数理论许多世纪的经典几何作图的情况一样。

最后的评论。我在这一节试图举例说明的主要哲学观点是,许多应用数学的基础可能获得一个简单的有穷论的表述和(如果愿意的话)一种有穷论的数值表征。在这个方向上更复杂的分析应用案例的某些努力,可以在我与丘瓦基(Rolando Chuaqui)合作的文章(1990,1995)中,或者,在我与理查德·萨默(Richard Sommer)合作的文章(1996,1997)中,还有参考文献里列出的苏佩斯和丘瓦基的文章(1993)中找到。

提出这种观点的一种不同方式是,在经验上考察已出版的大多数物理学文章和课本中用到的数学。文章中的大多数标准教诲或用法,在特征上是出奇地可作图的,而且,在实践中,正是有穷论的。很少能找到任何归纳证明,而且,几乎从来没有关于连续或收敛的 $\epsilon-\delta$ 证明。物理学家普遍使用的泰勒定理只要求展开式的头几项。关于物理学中经常出现的一些标准的微分方程的初步推导,参见苏佩斯和丘瓦基的文章(1993)。但是,对科学中的有穷论的数学方法的最有力的新的支持是,目前,在几乎所有的复杂问题中,都接近普遍地运用了有穷论的离散方法,因为它们在数字计算机上得到了实现,即在近期内似乎是不可能改变的一种局势。

7.
力学中的表征

本章我概述在力学中发挥突出作用的表征和不变性的一些 313 论题。第一节聚焦经典质点力学。本章作为经典物理的时空结构所介绍的实仿射空间概念,当然是6.1节介绍的经典时空的仿射空间的一种向量表述,这是一目了然的。第二节关注量子力学,但实际上只关注富有哲学意味的隐变量问题,并给出了关于它的各种不同的表征定理。值得强调的是,粒子的非定域的量子纠缠和相伴随的不存在经典的普通因果性(要不然称之为隐变量)是标准量子力学最持久的基本困惑。最后一节,即第三节,分析因果链的可逆性的两种重要但不同的意义。物理系统的时间可逆性是时间域的最重要的自然的不变性概念。对可逆性的含义的哲学反思通常对弱可逆性和强可逆性之间的基本区别,即在随机过程的理论中比在科学哲学中更熟悉的一种区别,始终不敏感。

7.1 经典质点力学

由于几种原因,我选择经典质点力学作为一个重要理论的另一个事例。首先,这个理论的直观内容是众所周知的。第二,

要求的一般数学框架相对简单和直接，但同时，根据它的哲学基础和它在真实世界中的系统应用，这个理论是丰富的，足以产生大量有意义的问题。

假定的数学概念。在完全转向力学之前，对在这些公理的陈述中和随后讨论这些公理的意义与含义时用到的某些数学概念，特别是几何概念，作出评论，将是令人向往的。首先，我很快评论前几章引入的一些概念。闭区间$[a, b]$是所有实数 x 的集合，使得 $a \leqslant x \leqslant b$。开区间$(a, b)$是所有实数 x 的集合，使得

314 $a < x < b$，相应地，$(a, b]$是一个半开、半闭区间，等等。区间$(-\infty, a)$是所有实数 x 的集合，使得 $x < a$，区间(a, ∞)是所有实数 x 的集合，使得 $a < x$，以及区间$(-\infty, \infty)$是所有实数的集合 Re。我们假设把关于实数及其特性的所有熟悉的运算都看成是已知的。

特别是，我们把实数域看成是已知的，它恰好是已知结构$(Re, +, \cdot, 0, 1)$，这里，$+$是实数的加法运算，$\cdot$ 是乘法运算，0 是数字零，1 是数字一。

我们现在转向实数域上的向量空间。一个向量空间具有集合论结构$(V, +, \cdot, \mathbf{0})$，这里，$V$ 是一个非空集，$+$是在 V 上的一种二元运算（向量加法），$\cdot$ 是从 $Re \times V$ 到 V 的一个函数（一个向量乘以一个实数产生另一个向量），$\mathbf{0}$ 是零向量。向量空间的公理在下列定义中给出。

定义 1. **一个结构 $\mathfrak{V} = (V, +, \cdot, \mathbf{0})$ 是 Re 上的一个向量空间，当且仅当，下列公理对 V 中的所有 $\mathbf{x}$、$\mathbf{y}$、$\mathbf{z}$ 和 Re 中的所有 a、b 都成立：**

1. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
2. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
3. $\mathbf{x} + \mathbf{0} = \mathbf{x}$

4. 在 V 中存在着 $\boldsymbol{u}$,使得 $\boldsymbol{x}+\boldsymbol{u}=\mathbf{0}$

5. $a(\boldsymbol{x}+\boldsymbol{y})=a\boldsymbol{x}+a\boldsymbol{y}$

6. $(a+b)\boldsymbol{x}=a\boldsymbol{x}+b\boldsymbol{x}$

7. $(ab)\boldsymbol{x}=a(b\boldsymbol{x})$

8. $1\boldsymbol{x}=\boldsymbol{x}$

公理 5－8 的符号是非正式的。以一种从初等算术开始就得到了广泛的使用的风格,用并置表示乘法,省略明显的符号"·"。受定义 3.1 的影响,对于群,很自然的是,引入作为基元的一元运算$-\boldsymbol{x}$,并把公理 4 写为 $\boldsymbol{x}+-\boldsymbol{x}=0$。鉴于向量空间的大多数标准的公理化,我们不这样做,但此后我们将把这个加法的逆符号看成是很明确的,并像所希望的那样用它。当然,我们为了把$-\boldsymbol{x}$也看成是明确的,还必须证明,公理 4 的向量 $\boldsymbol{u}$ 不仅存在,而且对于每个 $\boldsymbol{x}$ 都是惟一的。但这很容易做到。我们能够确立下列等式表明,我们不可能有 $\boldsymbol{u}\neq\boldsymbol{u}'$ 和 $\boldsymbol{x}+\boldsymbol{u}=\boldsymbol{x}+\boldsymbol{u}'=\mathbf{0}$。

$\boldsymbol{u}'=\mathbf{0}+\boldsymbol{u}'$	根据公理 2、3
$=(\boldsymbol{x}+\boldsymbol{u})+\boldsymbol{u}'$	根据关于 u 的假设
$=(\boldsymbol{x}+\boldsymbol{u}')+\boldsymbol{u}$	根据公理 1、2
$=\mathbf{0}+\boldsymbol{u}$	根据关于 $\boldsymbol{u}'$ 的假设
$=\boldsymbol{u}$	根据公理 2、3

重新表述公理 1－4 的一种简洁方式是,$(V,+)$是一个交换群,或者,在定义 2.3.1 的意义上,$(V,+,\mathbf{0},-)$是一个交换群。在线性代数中,定义 1 的更一般的形式通常是已知的。在一个任意域上,不只是在给定的实数域上,定义一个向量空间, 315 但为了力学的缘故,限于实数域是一种自然的限制。我们不说 $\mathfrak{V}$ 是 Re 上的一个向量空间,而通常是说,$\mathfrak{V}$ 是一个实向量空间。

一个向量空间最常见的模型是 Re^3，也就是，其中一个向量是实数的一个有序三元组的模型，即熟悉的解析几何的三维笛卡儿空间。第 6 章广泛地使用了这个笛卡儿模型，但也能使用这里采纳的更抽象的观点。然而我认为，笛卡儿模型越具体越好，尤其对视觉空间的广泛讨论来说，更是如此。在物理学中，向量的通常解释是，它们代表了像力之类的某个物理量相对于垂直的空间坐标轴的大小与方向。

我们以自然的方式定义向量的二元运算，或者，向量与实数的二元运算。因此，如果 $\boldsymbol{x} = (x_1, x_2, x_3)$ 是一个向量，a 是一个实数，那么

$$a\boldsymbol{x} = a(x_1, x_2, x_3) = (ax_1, ax_2, ax_3)$$
$$= (x_1a, x_2a, x_3a) = (x_1, x_2, x_3)a = \boldsymbol{x}a$$

如果 $\boldsymbol{x} = (x_1, x_2, x_3)$ 和 $\boldsymbol{y} = (y_1, y_2, y_3)$ 是向量，那么

$$\boldsymbol{x} + \boldsymbol{y} = (x_1, x_2, x_3) + (y_1, y_2, y_3)$$
$$= (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

和

$$-\boldsymbol{x} = -(x_1, x_2, x_3) = (-x_1, -x_2, -x_3)$$

同样，向量 $\mathbf{0} = (0, 0, 0)$。很容易检查，像刚才定义的那样，结构$(Re^3, +, \cdot, \mathbf{0})$是 Re 上的一个向量空间。

我们接下来定义两个向量的**标积**——也称为**内积**，它是从 $V \times V$ 到 Re 的一种映射或一个函数。任意两个向量的标积$(\boldsymbol{x}, \boldsymbol{y})$都有下列定义的特性：

1.（**双线性**）

$$(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z})$$
$$(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{x}, \boldsymbol{z})$$

$$(a\boldsymbol{x},\ \boldsymbol{y}) = a(\boldsymbol{x},\ \boldsymbol{y}) = (\boldsymbol{x},\ a\boldsymbol{y})$$

2.（对称性）

$$(\boldsymbol{x},\ \boldsymbol{y}) = (\boldsymbol{y},\ \boldsymbol{x})$$

3.（正定性）

如果 $\boldsymbol{x} \neq \boldsymbol{0}$，那么 $(\boldsymbol{x},\ \boldsymbol{x}) > 0$。

此外，**欧几里得范数** $|\boldsymbol{x}|$ 是非负值的函数，定义如下：

$$|\boldsymbol{x}| = (\boldsymbol{x},\ \boldsymbol{x})^{\frac{1}{2}}$$

而且，**欧几里得度量** $d(\boldsymbol{x},\ \boldsymbol{y})$ 是非负的函数，定义为

$$d(\boldsymbol{x},\ \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}| = (\boldsymbol{x} - \boldsymbol{y},\ \boldsymbol{x} - \boldsymbol{y})^{\frac{1}{2}}$$

标积具有几何解释：

$$(\boldsymbol{x},\ \boldsymbol{y}) = |\boldsymbol{x}||\boldsymbol{y}|\cos(\boldsymbol{x},\ \boldsymbol{y})$$

这里，$\cos(\boldsymbol{x},\ \boldsymbol{y})$ 是向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 夹角的余弦。在已经讨论过的一个向量空间的笛卡儿模型中，标积是 316

$$((x_1,\ x_2,\ x_3),\ (y_1,\ y_2,\ y_3)) = x_1 y_1 + x_2 y_2 + x_3 y_3$$

为了以简单的向量形式表示牛顿第三定律，我们也需要陈述两个向量的**矢积**——也称为**外积**——的定义特性，它是从 $V \times V$ 映射到 V。任意两个向量的**矢积** $[\boldsymbol{x},\ \boldsymbol{y}]$ 具有下列定义特性：

1.（双线性）

$$[\boldsymbol{x} + \boldsymbol{y},\ \boldsymbol{z}] = [\boldsymbol{x},\ \boldsymbol{z}] + [\boldsymbol{y},\ \boldsymbol{z}]$$

$$[\boldsymbol{x},\ \boldsymbol{y} + \boldsymbol{z}] = [\boldsymbol{x},\ \boldsymbol{y}] + [\boldsymbol{x},\ \boldsymbol{z}]$$

$$[a\boldsymbol{x},\ \boldsymbol{y}] = a[\boldsymbol{x},\ \boldsymbol{y}] = [\boldsymbol{x},\ a\boldsymbol{y}]$$

2.（斜对称）

$$[\boldsymbol{x},\ \boldsymbol{y}] = -[\boldsymbol{y},\ \boldsymbol{x}]$$

3. (雅可比等式)

$$[[\boldsymbol{x},\ \boldsymbol{y}],\ \boldsymbol{z}]+[[\boldsymbol{y},\ \boldsymbol{z}],\ \boldsymbol{x}]+[[\boldsymbol{z},\ \boldsymbol{x}],\ \boldsymbol{y}]=\mathbf{0}$$

4. (正交性)

$$[[\boldsymbol{x},\ \boldsymbol{y}],\ \boldsymbol{x}]=\mathbf{0}$$

在前面讨论过的一个向量空间的笛卡儿模型中,矢积具有下列特征描述:

$$[(x_1,\ x_2,\ x_3),\ (y_1,\ y_2,\ y_3)]$$
$$=(x_2y_3-x_3y_2,\ x_3y_1-x_1y_3,\ x_1y_2-x_2y_1)$$

有些标准的向量概念也是我们在本节中需要的。首先,向量 $\boldsymbol{x}_1,\ \cdots,\ \boldsymbol{x}_n\in V$ **是线性独立的**,当且仅当,对于任意实数 $a_1,\ \cdots,\ a_n$,如果 $\sum_n a_i\boldsymbol{x}_i=0$,那么 $\boldsymbol{x}_i=0$。向量空间 $\mathfrak{V}$ 在 Re 上有 n 维,当且仅当,n 是 $\mathfrak{V}$ 的线性独立向量的最大数。如果 $\mathfrak{V}$ 有 n 维,那么 $\mathfrak{V}$ 的一个基是 $\mathfrak{V}$ 的 n 个线性独立向量的任何一个集合。尽管经典物理学的空间是 3 维,但在力学中,需要更大维度的向量空间。

我们需要的下一个概念是表示质点位置的一个仿射空间概念。在经典力学的较早的公理化中(McKinsey, Sugar and Suppes 1953; Suppes 1957/1999),我们用 Re^3 中的一个向量表示位置,但更可取的是与坐标无关的表征。在这方面,向量空间已经是一种改进,但因为一个向量空间有一个自然原点,这也是不能完全令人满意的。一个任意的仿射空间似乎是表征经典空间和时间因而表征质点位置的适当框架。①

---

① 正如读者很容易确定的那样,在首先从综合几何的观点描述仿射空间(即实际上只用作为基元的介中性的三元关系)的地方,这种分析与 6.1 节的分析密切相关。

有几种不同的但本质上等价的方式，把一个实仿射空间概念(即与实数域上的向量空间相联系的一个仿射空间概念)公理化。设 A 是我们赋予一个仿射结构的一个任意的非空集。A 的元素我们称为**点**。如此赋予 A 的一种方式是，通过在 A 上假定的实向量空间 $\mathfrak{V}$ 的**作用**(action)，更精确地说，通过在 A 上的 $\mathfrak{V}$ 的加法群的作用。这样一种作用是从 $V \times A$ 到 A 的一个函数，我们也用加法符号来表示，而且，它满足下列三个公理： 317

(1) 如果 $\boldsymbol{x}$、$\boldsymbol{y}$ 在 V 中和 P 在 A 中，那么

$$(\boldsymbol{x}+\boldsymbol{y})+P=\boldsymbol{x}+(\boldsymbol{y}+P)$$

(2) $\mathbf{0}+P=P$，其中，$\mathbf{0}$ 是在 V 中的零向量。

(3) 对于 A 的每个有序点对$(P,\ Q)$，恰好存在着在 V 中的一个向量 $\boldsymbol{x}$，使得

$$\boldsymbol{x}+P=Q$$

特别是由于最后这个方程，V 的加法群通常被称为 V 的**平行位移的群**。

尽管这条进路及其相联系的术语很流行(例如，参见，Snapper & Troyer 1971, and Arnold 1978)，但似乎是不自然的，尤其在符号上不自然，因为用“+”表示把不同类型的对象结合起来的一种二元运算，是不常见的，而且，用两点 P 和 Q 不能定义加法。一条稍微不同的进路强调，关于点的有意义的运算是差运算，即从 $A\times A$ 映射到 V。习惯上不把差写成 $Q-P$，而是写成$\underline{PQ}$(或者有时在“PQ”的上面加箭头，表明向量是从 P 指向 Q)。这里将用符号 $Q-P$ 和$\underline{PQ}$。无论如何，我在后面的表述中主要用到这种减法，但根据差，很容易把相加的位移运算定义为

$$\boldsymbol{x}+P=Q\text{，当且仅当，}Q-P=\boldsymbol{x}$$

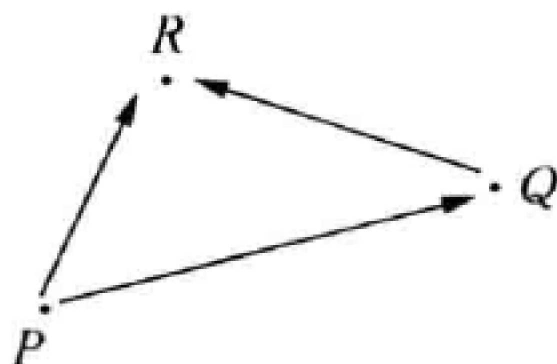
定义 2. 一个结构 $\mathfrak{A}=(A, \mathfrak{B})$ 是一个实仿射空间，当且仅当，A 是一个非空集，$\mathfrak{B}$ 是一个实向量空间，并满足下列公理：

1. 对于 A 中的任意点 P 和 $\mathfrak{B}$ 中的向量 $\boldsymbol{x}$，在 A 中恰好有一点 Q，使得 $\underline{PQ}=\boldsymbol{x}$。

2. 对于 A 中的任意点 P、Q 和 R，

$$\underline{PQ}+\underline{QR}=\underline{PR}$$ ①

对于这些点来说，第二个公理是更重要的向量加法定律。



证明关于实仿射空间的一些基本事实，将是有用的。首先，在公理 2 中，我们指定 $Q=P$ 并得到

$$\underline{PP}+\underline{PR}=\underline{PR}$$

318 因而对于 A 中的每个点 P，$\underline{PP}=\mathbf{0}$——注意，$\mathbf{0}$ 是 $\mathfrak{B}$ 的零向量。现在，在公理 2 中指定 $R=P$，我们有

$$\underline{PQ}+\underline{QP}=\underline{PP}$$

因此，

$$\underline{PQ}=-\underline{QP}$$

我们留给读者证明，$\underline{P_1Q_1}=\underline{P_2Q_2}$ 意味着 $\underline{P_1P_2}=\underline{Q_1Q_2}$，这是一种形式的平行四边形法则。实仿射空间$(A, \mathfrak{B})$的维度恰好

① 比较仿射空间的这个定义与 6.1 节中的定义。在这个定义中，即在一个实向量空间，有更多的假定。在满足这两个不同定义的那些结构之间如何构造出适当的同构性，我留给读者来完成。

是实向量空间 $\mathfrak{V}$ 的维度。为了区分维数，我们有时根据三维向量空间 $\mathfrak{V}(3)$ 通常是四维向量空间 $\mathfrak{V}(4)$ 的一个子空间，写 $\mathfrak{V}(3)$ 或 $\mathfrak{V}(4)$。当 $\mathfrak{V}$ 有前面所定义的一个内积和欧几里得范数时，$(A, \mathfrak{V})$ 被说成是一个**欧几里得空间**，并且，两点 P 和 Q 之间的**距离** $d(P, Q)$ 恰好是向量 $\underline{PQ}$ 的欧几里得范数，即

$$d(P, Q) = |\underline{PQ}| = (\underline{PQ}, \underline{PQ})^{\frac{1}{2}}$$

很容易表明，点之间的距离具有一个度量空间的特性：

1. $d(P, Q) \geqslant 0$
2. $d(P, Q) = 0$，当且仅当，$P = Q$
3. $d(P, Q) = d(Q, P)$
4. $d(P, Q) \leqslant d(P, R) + d(R, Q)$

时空结构。接下来定义的重要结构是经典时空结构。

定义 3. 设 $\mathfrak{A} = (A, \mathfrak{V})$ 是一个四维的实仿射空间，设 τ 是从 A 到 Re 的一种映射，再设 $\boldsymbol{\alpha}$ 是 τ 的值域到 A 的一种映射。于是，$(\mathfrak{A}, \tau, \boldsymbol{\alpha})$ 是一个经典时空结构，当且仅当，对于 A 中的每个点 P，$\{Q \mid \tau(Q) - \tau(P) = 0\}$ 是前面在三维实向量空间 $\mathfrak{V}(3)$ 上定义的距离函数的三维欧几里得空间，并且，$\boldsymbol{\alpha}(\tau(P))$ 是这个三维欧几里得空间中的一个点，即

$$\tau(\boldsymbol{\alpha}(\tau(P))) - \tau(P) = \mathbf{0}$$

应该显而易见的是，映射 τ 是计时时间函数(clock time-function)，它给出了 A 中的任何一个时空点事件发生的时间。集合 $\{Q \mid \tau(Q) - \tau(P) = 0\}$ 是在 A 中与 P 同时发生的所有点(或更生动地说是点事件)的集合。对于在 τ 值域内的每个 t 来说，我们用符号 $\mathfrak{A}(t)$ 表示在时间 t 发生的所有点事件组成的 $\mathfrak{A}$ 的三维仿射子空间。从定义 3 立即得出，$\mathfrak{A}(t)$ 是欧几里得空间。注意，当我们考虑经典结构的表征定理时，$\mathfrak{A}$ 本身不是欧几里得

空间，只是仿射空间，因为比较不同时发生的两个事件之间的距离，不可能赋予任何客观意义，这是我们上一章更详细地探讨的一个问题。在当前的语境中，这意味着与四维仿射空间相联系的向量空间没有内积。

319 从 τ 的值域，即瞬时测量，到 A 的函数 $\boldsymbol{\alpha}$，具有明显的和重要的物理意义。对于每个瞬时 t 来说，A 的点 $\boldsymbol{\alpha}(t)$ 是运动所参照的固定的惯性参照系的原点。通过把 $\boldsymbol{\alpha}$ 推广到所有惯性参照系的族，能够给出一种更加不变的表述。在赫米斯(Hermes, 1938)的文章和苏佩斯(Suppes, 1959)的著作中，显然这是为狭义相对论做的，但就当前的目标而言，这种做法引入了太多的注解，并使我们过分远离了本节后面集中讨论的力学的传统符号。

关于 $\boldsymbol{\alpha}$ 的形式要求的意义是明显的。惯性参照系的原点必须随时存在。

我们在陈述经典力学的公理时假定，从一个实数区间 T 到已知内积的一个实向量空间，一个位置函数概念 $\mathbf{s}$，在 T 的每一点，都是可微的或二次可微的，因此在 T 上。同样，我们假定，当然的意思是指，把 $\mathbf{s}$ 关于 T 的一阶导数和二次导数写成 $\frac{d}{dt}\mathbf{s}$ 和 $\frac{d^2}{dt^2}\mathbf{s}$，或写成 $D\mathbf{s}$ 和 $D^2\mathbf{s}$，或写成 $\dot{\mathbf{s}}$ 和 $\ddot{\mathbf{s}}$。(在下面熟悉的数学和物理学实践中，将不止用一种导数符号；当用符号 D 或 D^2 时，微分变量是明显的。)

不难表明，通常的微分乘法定则对于标积和矢积都成立。为了明确起见，设 $\boldsymbol{x}(t)$ 和 $\boldsymbol{y}(t)$ 是可微的向量值函数。于是，

$$\text{标积：}\ \frac{d}{dt}(\boldsymbol{x}(t),\ \boldsymbol{y}(t))=\left(\frac{d}{dt}\boldsymbol{x}(t),\ \boldsymbol{y}(t)\right)+\left(\boldsymbol{x}(t),\ \frac{d}{dt}\boldsymbol{y}(t)\right)$$

和

矢积：$\frac{d}{dt}[\boldsymbol{x}(t), \boldsymbol{y}(t)]=\left[\frac{d}{dt}\boldsymbol{x}(t), \boldsymbol{y}(t)\right]+\left[\boldsymbol{x}(t), \frac{d}{dt}\boldsymbol{y}(t)\right]$

对于可数的向量之和,即向量的一个无穷级数,我们也需要绝对收敛的概念。这种要求恰好是,数值级数 $\sum_{i=1}^{\infty}|\boldsymbol{x}_i|$ 是绝对收敛的,其中,$|\boldsymbol{x}_i|$是欧几里得范数。

原始概念。我们下一步的任务是,描述这里给出的力学公理化的原始概念。这种公理化是相对于一个已知的经典时空结构的。设 T 是 τ 的值域。我们用了五个原始概念：质点集合 P;从笛卡儿积 $P\times T$ 到 A(即已知的仿射空间)的位置函数 $\boldsymbol{s}$;从质点集合 P 到 Re 的质量函数 m;代表内力的一个力函数 $\boldsymbol{f}$,即从质点集合本身和时间区间 T 的笛卡儿积,即 $P\times P\times T$ 到三维向量空间 $\mathfrak{V}(3)$的函数;以及从质点集合、集合 T 和正整数集合的笛卡儿积,即 $P\times T\times \underline{I}^{+}$ 到 $\mathfrak{V}(3)$的外力函数 $\boldsymbol{g}$。为了提供计算外力的一种方法,外力函数的定义中包括正整数。

为了使这些概念变得完全明确,我以稍微不同的方式陈述它们应有的解释。如果 $p\in P$ 和 $t\in T$,那么 $\boldsymbol{s}(p, t)$在物理上被解释为质点 p 在时间 t 的位置。当然,所观察到的是矢量 $\boldsymbol{s}(p, t)-\boldsymbol{\alpha}(t)$,而不是无条件的 $\boldsymbol{s}(p, t)$。对于 P 中的每个 p, 320 我们通常把它的位置函数写成 $\boldsymbol{s}_p$。对于 P 中的 p,$m(p)$被解释为质点 p 的质量的数值。如果 p 和 q 在 P 中,t 在 T 中,那么 $\boldsymbol{f}(p, q, t)$是质点 q 在时间 t 施加于质点 p 的力。这样的力是力学系统的内力。如果 p 在 P 中,t 在 T 中,n 是正整数,那么 $\boldsymbol{g}(p, t, n)$是在时间 t 作用于质点 p 的第 n 个外力。

在描述力学的原始概念时,我们用了习惯性的假设：原始函数和关系不是在比这些公理中提到的那些集合更大的任何一个集合上下定义。例如,只有当 p 在 P 中,才能定义 $m(p)$;同

样,只有当 p 在 P 中和 t 在 T 中等等时,才能定义 $\boldsymbol{s}(p, t)$。从这些考虑得出的结论是,基元 P 实际上根据质量函数 m 的定义域是可定义的。然而,在公理的通常的数学表述中也习惯于不排除这样的基元,恰好因为它们是可定义的,因此,我们不坚持根据独立的原始概念集来表述这些公理。

公理。我们现在准备表述质点力学系统的公理。这些公理只假定已经明确讨论过的那些数学概念,不是物理学的其他部分。因此,这些公理确实提供了谓词"是一个经典质点力学系统"的纯集合论定义的核心。沿着经典思路,把这些公理本身划分为运动学的公理和动力学的公理。支持这种区分的直观思想是,运动学的公理是只要求描述质点运动的那些公理。动力学的公理纳入到关于运动原因的任何考虑之中。我们在明确地陈述了这些公理之后,进一步详述这种区分。

定义 4. 一个结构 $\mathfrak{P}=(P, \boldsymbol{s}, m, \boldsymbol{f}, \boldsymbol{g})$ 是一个关于经典时空结构$(\mathfrak{A}, \tau, \boldsymbol{\alpha})$的(经典的)质点力学系统,当且仅当,满足下列公理:

运动学公理

1. 集合 T,即 τ 的值域,是一个实数区间。

2. 集合 T 是有穷的和非空的。

3. 对于在 P 中的每个 p 和在 T 中的 t,$s_p(t)$在 $\mathfrak{A}(t)$中。

4. 对于在 P 中的每个 p 和在 T 中的 t,向量函数 $s_p(t)-\boldsymbol{\alpha}(t)$ 是在 t 的二次微分。

动力学公理

5. 对于 P 中的 p,$m(p)$是正实数。

6. 对于 P 中的 p 与 q 以及 T 中的 t,

$$\boldsymbol{f}(p, q, t)=-\boldsymbol{f}(q, p, t)$$

7. 对于 P 中的 p 与 q 以及 T 中的 t，

$$[\mathbf{s}(p, t) - \mathbf{s}(q, t), \mathbf{f}(p, q, t) - \mathbf{f}(q, p, t)] = \mathbf{0}$$

8. 对于 P 中的 p 和 T 中的 t，级数， 321

$$\sum_{n=1}^{\infty} \mathbf{g}(p, t, n) \text{ 是绝对收敛的。}$$

9. 对于 P 中的 p 和 T 中的 t，

$$m(p)D^2(\mathbf{s}_p(t) - \boldsymbol{\alpha}(t)) = \sum_{q \in P} \mathbf{f}(p, q, t) + \sum_{n=1}^{\infty} \mathbf{g}(p, t, n)$$

与坐标无关的数学表述和仿射空间 $\mathfrak{A}$ 的明确引入是上述公理与在麦金西(McKinsey)、休格(Sugar)和苏佩斯(1953)以及苏佩斯(1957/1999)的文章中发现的那些公理之间的主要差别。仿射空间的明确引入使这些公理变得更加复杂，但依我的判断，这却决定性地完善了在使力学的物理内容变得更加完全明确时所提到的早期的数学表述。在早期的数学表述中，从一开始就假设了笛卡儿坐标系，而且，空间的仿射特征不明显。现在，似乎特别重要的是引入参照点 $\boldsymbol{\alpha}(t)$，即"固定参照系"的原点。①绝对位置在仿射空间 $\mathfrak{A}(t)$ 中没有经验意义，只是相对位置，这是经验事实。第二，通过在仿射空间中引进位置和在相联系的向量空间中引入力，强调了这种物理差别，早期的工作并非如此。

关于这些公理有几种更直观的概念评论。我首先评论运动学的公理。在公理 1 的情况下，我们愿意把任何一个实数区间看成是时间测量。这个区间是整个实数集合，对于许多应用，事实上，是大多数实际应用来说，这肯定是不必要的。我们也应该

---

① 从直观的和非形式的观点来看，在阿佩尔(Appell, Tome I, P100)的经典著作中，通过假定 $\boldsymbol{\alpha}(t)$ 是太阳系的质心，非常漂亮地做到了这一点。

注意到，在这个公理中，我们已经看到了引入数学便利的成分，这与关于经验约束的任何问题完全无关。站在经验的立场上，将一定不可能区分公理1和下列公理：在理论上可取的整个实数区间内，流逝的时间的集合 T 是代数的集合乃至是有理数的集合，尽管这种变化要求修改其他公理，特别是公理9。整个实数区间，对于达到应用数学分析的目的，尤其是达到应用标准的微积分方法和微分方程的目的都是有用的，但重要的是承认，这种方法根据它产生的计算结果类型被证明是合理的，而不是因为能够直接根据可观察的经验数据，乃至源于经验数据的任何相当间接的论证，来证实位置函数的可微性。

公理2中表达的必要条件：即质点集合是有穷的，与质点力学的通常阐述相一致。如果我们拒绝接受这个必要条件，那么为了使公理9中的运动第二定律的表达式令人满意，有必要322 为内力增加某种收敛条件。此外，我们可能希望增加的附加约束是，要求系统的总质量，也许还有系统的动能，是有穷的。①

关于公理4应该注意，稍弱的必要条件是通过哈梅尔

① 概括地说，从假定质点的有穷集转到无穷集是连续介质力学的典型特征。因此，正如所预期的那样，这些公理是非常复杂的，通过研究沃尔特·诺尔(Walter Noll)的重要文章能看出这一点，诺尔的这篇论文于1974年再版。诺尔是对近代连续介质力学基础作出杰出贡献的人之一。当然，也有一种重要的可数的无穷情况，即理想化的统计力学版本。关于连续介质力学史和统计力学史的许多出色的远见卓识，我极力推荐参考特鲁斯德尔(Truesdell)的著作(1968)，因为它在风格上很活泼且有直率的见解。从纯认识论的观点来看，我强烈地认为，质点数总是有穷的，但如果这个数只是有穷的话，它是如此之大，以至于假设无穷的渐近方法，完全是为了在实践中获得具体结果。但很容易陈述的我的观点，在某种意义上，代表了具有卓越谱系的一种哲学偏见。统治了近代物理学的许多方面的物质的场进路，没有使物质的质点观处于最显著的重要地位。诺尔(1955)表明，在相当普遍的条件下，统计力学的平均量——比如，平均速度、平均能量流——满足连续物质的场方程。因此，就最重要的物理量而言，我们无法从理论上在流体或固体的质点或分子的观点和场的观点之间作出区分。

(Hamel, 1958)的一个公理强加的,该公理只要求,位置函数是分段二次可微的(一个函数是分段可微的,如果除了有穷数目的点之外,它在定义它的任何一个有穷区间的所有点上都是可微的)。对于碰撞作用的理想化分析来说,这种分段限制是便利的。例如,合理的近似是假定,当两个刚体碰撞时,在碰撞点,它们的位置函数是不可微的,尽管这些函数是连续的。为了达到当前讨论的概念目标,哈梅尔公理的这种稍微更现实的假设引入了不需要探讨的复杂性,因此,我们还是使用更受限制的和更简单的公理 4。

正如已经表明的那样,前四个公理构成质点力学系统的运动学公理。显然,这些公理的一般形式是相当弱的,但这将强调,质点问题可能是格外困难的。也许,最漂亮的两个经典事例是托勒密的对行星运动的分析和开普勒关于行星运动的三个运动学定律的推导。托勒密的《天文学大成》中关于行星运动的详细结果代表了古代经验科学最有意义的工作。古代天文学拒绝接受希腊的科学遗产在特征上是非经验的主张。无论如何,我提到这些历史上重要的运动学结果来强调,几个世纪的力学发展蕴含了比动力学或因果律的发现更多的东西。如果在牛顿之前,没有从二千多年的持久的天文学观察中抽象出的运动学规律性的话,那么牛顿在表述和解决两体引力问题(他的最伟大的成就)时会如此成功,是值得怀疑的。①

第一个动力学公理(即公理 5)断言了标准必要条件:一个质点的质量是正的。在文献中,对这个公理有各种不同的概括, 323

① 牛顿关于实际的有穷物体在流体中运动的结果,即连续介质或流体力学的一个重要部分,没有给人留下多少深刻的印象,这在某种程度上是因为早期没有大量发展纯粹描述的、非因果性的理论。不管怎样,关于牛顿在流体力学方面的工作的好的讨论和评价,参见 Truesdell(1968, 第 3 章)。

例如,允许质量是零或负的概括,或者,允许质量随时变化的概括,但这里考察这些技术变化似乎与主题无关。

公理6和公理7共同对牛顿第三运动定律提供了一个精确的表述。公理6对应于哈梅尔(1958,第25页)所称的第一完全反作用原理。这个原理要求,质点 q 施加于质点 p 的力与质点 p 施加于质点 q 的力,在数值上大小相等,而方向相反。公理7(哈梅尔称之为第二完全反作用原理)要求,力的作用线平行于两个质点位置的连线。(我们作为一种练习留给读者证明,从矢积的定义中得出的结论是,公理7确实要求,力的作用线平行于质点 p 和 q 的连线。)

公理8要求,作用于一个质点的外力之和是绝对收敛的。要求绝对收敛而不是简单收敛的理由是,质点的运动应该与命名适用于收敛的质点的顺序无关。在当前的语境中,绝对收敛是具有很强物理学基础的一个对称性原理或不变性原理。对作用于一个质点的外力进行编号所确定的任意顺序,没有明显的物理意义。

最后,公理9是在牛顿第二定律的当前语境中的数学表述。在考虑质点力学的基础时,经常会讨论两个另外的公理。第一个公理只是牛顿第一定律的表达。作为一个公理,我省略了这一点,因为它是当前公理的一个不重要的结论(参见下面的定理1)。第二个事例是一个不可切入公理(axiom of impenetrability),这个公理表述如下。对于 P 的任何两个不同的质点 p 和 q 以及 T 中的任意时间 t,$s(p, t)$不等于 $s(q, t)$。这样一个不可切入公理被省略了,主要因为质点力学系统的标准解释之一是,这些质点是刚体的质心,而且,很容易想到这方面的物理学事例,例如,当球抛出去穿过一个圆环时,当子弹打出去穿越一个圆洞时,当一个人坐在飞机上的某个地方时,这些不同物体的质心在

某个时刻可以相一致。另一方面，我们应该注意到，在 18 世纪欧拉等人讨论物质理论时，这个不可切入公理起着重要的作用。①

两个定理——有一个定理是关于决定论的。如前所述，第一个定理是牛顿第一定律的表述。在这个定理和后面的定理中，当然力学系统与一个给定的经典时空结构(A, τ, α)相关。

定理 1. *设 $\mathfrak{P} = (P, \mathbf{s}, m, \boldsymbol{f}, \boldsymbol{g})$ 是一个质点力学系统，设 p 是 P 的项，使得对于 T 中的所有 t，即 τ 的值域，* 324

$$\sum_{q \in P} \boldsymbol{f}(p, q, t) + \sum_{n=1}^{\infty} \boldsymbol{g}(p, t, n) = 0$$

于是，存在向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$，使得对于 T 中的所有 t，

$$\mathbf{s}(p, t) - \boldsymbol{\alpha}(t) = \boldsymbol{a} + \boldsymbol{b}t$$

我们立刻从公理 4、5 和 9 得到这个定理的证明。

下一个定理是下列事实的经典表达：即质点力学是一门完全决定论的科学。这个定理说，一个系统的整个历史是由 P、T、m、$\boldsymbol{f}$、$\boldsymbol{g}$ 和适当的初始条件决定的。特别是，如果我们知道作用于一个质点集合的质量和力，还有它们在一个已知时刻的位置与速度，我们就知道了这个系统的全部历史。证明取决于对下列问题的基本考虑：即由牛顿第二定律例示的这种微分方程的解具有惟一性，这里省略了证明。

定理 2. *设 $\mathfrak{P} = (P, \mathbf{s}, m, \boldsymbol{f}, \boldsymbol{g})$ 和 $\mathfrak{P}' = (P, \mathbf{s}', m, \boldsymbol{f}, \boldsymbol{g})$ 是两个质点力学系统，使得对于 T 中的某个 t 和 P 中的每个 p，*

$$\mathbf{s}(p, t) - \boldsymbol{\alpha}(t) = \mathbf{s}'(p, t) - \boldsymbol{\alpha}(t)$$

---

① 欧拉在许多不同的地方讨论了不可切入性，但也许他的最好的哲学讨论是他在 1842 年就物理学和哲学各种问题写给德国公主的一系列信件的第二部分的头两封信中，其中，他也批评了笛卡儿作为纯外延的物质理论。

和

$$D(\boldsymbol{s}_p(t) - \boldsymbol{\alpha}(t)) = D(\boldsymbol{s}'_p(t) - \boldsymbol{\alpha}(t))$$

于是，对于 T 中的每个 t 和 P 中的每个 p，

$$\boldsymbol{s}(p, t) - \boldsymbol{\alpha}(t) = \boldsymbol{s}'(p, t) - \boldsymbol{\alpha}(t)$$

因为一个质点的轨迹或轨道完全是由它受到的作用力和某个时间 t 的位置与速度的初始条件决定的，所以，我们易于错误地以为，我们能够经常直接求解运动的速度方程——笛卡儿坐标中的三个方程——获得作为 t 的**显函数**的质点轨迹，这个显函数是由数目有穷的已知数学函数组成的。不幸的是，事实远非如此。甚至对于有一个简单表述因而容易推导运动方程的物理情境来说，刚才描述的那类解也可能没有闭合解。最著名的事例是三体问题：确定只受到彼此引力作用的三个质点的运动。同其他情况一样，在这种情况下，用幂级数表示解总是可能的，但这样的无穷级数，对于研究一个系统的定性行为或长期行为来说，根本不会令人满意。就 5.6 节中的一个重要的特例而言，三体问题的复杂性已经体现出来。[①] 在各种特殊情况的解中运用的独创性是力学的荣耀，其中，最著名的历史事例是牛顿关于两体引力问题的解。

325 **动量和角动量**。本节关于经典力学的其余部分，定位于关于动量和角动量的某些一般定理，这些定理也要求明显地引入陈述这些公理不需要的最有意义的力学概念。

为了用一个更简洁的符号表示一个质点在时间 t 的位置向量，我们将用熟悉的符号 $\boldsymbol{q}_i(t)$，在这种情况下，现在我们也用数

---

① 近来，在三体问题的某些新的特殊解方面取得了惊人的进步，对这个问题的具有可读性的说明，参见 Montgomery(2001)。

字索引 i 代替质点 p 的"p",因此,

$$\boldsymbol{q}_i(t) = \boldsymbol{s}(i,\ t) - \boldsymbol{\alpha}(t)$$

在这里,为了与标准的物理学符号更加一致,我们写 $\boldsymbol{q}_i$,而不是写 $\boldsymbol{q}_i(t)$,这不会造成混淆。我们也需要一个作用于质点 i 的合力的(resultant)符号,其中,N 是质点数:

$$\boldsymbol{F}_i(t) = \sum_{j=1}^{N} \boldsymbol{f}(i,\ j,\ t) + \sum_{n=1}^{\infty} \boldsymbol{g}(i,\ t,\ n)$$

同样,我们通常写 $\boldsymbol{F}_i$,而不是写 $\boldsymbol{F}_i(t)$。第二,我们需要一个作用于质点 i 的合外力的符号:

$$\boldsymbol{G}_i(t) = \sum_{n=1}^{\infty} \boldsymbol{g}(i,\ t,\ n)$$

质点 i 的(线性)动量是 $m_i\boldsymbol{q}_i$,从这个定义得出,**一个质点的动量的变化率等于作用于这个质点的合力**。用符号表示为

$$\frac{d}{dt}(m_i\, \dot{\boldsymbol{q}}_i) = \boldsymbol{F}_i$$

在这种形式中,该定律通常指**线性动量原理**。把该原理应用于一个系统,产生了一个有意义的结果。首先,我们把一个质点系统的动量 $\boldsymbol{P}(t)$ 定义为

$$\boldsymbol{P}(t) = \sum_{i=1}^{N} m_i\, \dot{\boldsymbol{q}}_i(t)$$

其中,N 是质点数。

定理 3. *一个质点系统的动量的变化率等于作用于这些质点的合外力之和*,即

$$\frac{d}{dt}\boldsymbol{P}(t) = \sum_{i=1}^{N} \boldsymbol{G}_i(t)$$

证明很简单。它尤其依赖于公理 6,即内力大小相等方向

相反的原理,这是在确定 $\boldsymbol{P}(t)$ 的变化率时取消内力的基础。

我们现在把定理 3 推广到一个系统的质心。设

$$\underline{m} = \sum_{i=1}^{N} m_i$$

即设 $\underline{m}$ 是该系统的总质量。一个系统的质心 $\boldsymbol{c}(t)$ 定义如下。

326 设 $\boldsymbol{b}(t)$ 是在 $\mathfrak{A}(t)$ 中的任何一个仿射点。关于质点系统 $\mathfrak{P}$ 的 $\boldsymbol{b}(t)$ 的(**线性**)**力矩** $\boldsymbol{L}(t)$ 定义为

$$\boldsymbol{L}(t) = \sum_{i=1}^{N} m_i(\boldsymbol{s}_i(t) - \boldsymbol{b}(t))$$

于是,$\mathfrak{P}$ 的**质心**定义为点 $\boldsymbol{c}(t)$,使得关于 $\mathfrak{P}$ 的 $\boldsymbol{c}(t)$ 的(线性)力矩为零。当然,为了使这个定义成为严格的定义,我们必须证明,这样一个点 $c(t)$ 是存在的,而且是惟一的。考虑向量

$$\frac{\sum m_i(\boldsymbol{s}_i(t) - \boldsymbol{\alpha}(t))}{\underline{m}}$$

和仿射点 $\boldsymbol{\alpha}(t)$。根据仿射空间的定义 2,存在惟一的点 $\boldsymbol{c}(t)$,使得

$$\frac{\sum m_i(\boldsymbol{s}_i(t) - \boldsymbol{\alpha}(t))}{\underline{m}} = \boldsymbol{c}(t) - \boldsymbol{\alpha}(t)$$

因此,

$$\sum m_i(\boldsymbol{s}_i(t) - \boldsymbol{\alpha}(t)) = \underline{m}(\boldsymbol{c}(t) - \boldsymbol{\alpha}(t))$$

那么

$$\sum m_i(\boldsymbol{s}_i(t) - \boldsymbol{c}(t)) = \underline{m}(\boldsymbol{\alpha}(t) - \boldsymbol{\alpha}(t)) = 0 \tag{1}$$

既然对于三个仿射点 P、Q 和 R,

$$\underline{PQ} + \underline{QR} = \underline{PR}$$

即

$$\underline{QR} = \underline{PR} - \underline{PQ}$$

或者用不同的符号表示为

$$Q - R = (P - R) - (P - Q) \tag{2}$$

我们从(1)和(2)推出

$$\sum m_i(\boldsymbol{s}_i(t) - \boldsymbol{c}(t)) = 0$$

因此,所定义的质心总是存在的,并且是惟一的。

我们留给读者证明下列经典表征。

定理 4. 一个质点系统的质心像一个质点一样运动,这个质点具有的质量等于该系统的总质量,这个质点所受的作用力等于该系统的合外力。

我们现在转向角动量。一个质点相对于点 $\boldsymbol{b}(t)$ 的角动量 $\boldsymbol{M}(t)$ 是相对于 $\boldsymbol{b}(t)$ 的(线性)动量矩,即质点 i 相对于 $\boldsymbol{b}(t)$ 及其动量的位置向量的矢积。用符号表示为

$$\boldsymbol{M}_i(t) = [\boldsymbol{s}_i(t) - \boldsymbol{b}(t),\ m_i D(\boldsymbol{s}_i(t) - \boldsymbol{b}(t))]$$

在选择 $\boldsymbol{b}(t)$ 为 $\boldsymbol{\alpha}(t)$ 的情况下,我们有熟悉的公式

$$\boldsymbol{M}_i(t) = [\boldsymbol{q}_i,\ m_i \dot{\boldsymbol{q}}_i]$$

相对于一点的质点系统的角动量恰好是

$$\boldsymbol{M}(t) = \sum_{i=1}^{N} \boldsymbol{M}_i(t)$$

作用于关于点 $\boldsymbol{b}(t)$ 的质点 i 的力的**力矩**或**转矩**正好是　　327

$$[\boldsymbol{s}(i,\ t) - \boldsymbol{b}(t),\ \boldsymbol{F}]$$

定理 5. 关于一点的质点系统的角动量的变化率等于作用于该系统的质点的外力的力矩之和,用符号表示为

$$\frac{d\boldsymbol{M}(t)}{dt} = \sum_{i=1}^{N} [\boldsymbol{s}(i,\ t) - \boldsymbol{b}(t),\ \boldsymbol{G}_i]$$

证明： 根据 $\boldsymbol{M}(t)$的定义

$$\frac{d\boldsymbol{M}(t)}{dt} = \sum_{i=1}^{N} ([D(\boldsymbol{s}_i(t) - \boldsymbol{b}(t)),\ m_i D(\boldsymbol{s}_i(t) - \boldsymbol{b}(t))]$$
$$+ [\boldsymbol{s}(t) - \boldsymbol{b}(t),\ m_i D^2(\boldsymbol{s}_i(t) - \boldsymbol{b}(t))])$$

第一个矢积是零,第二个矢积的第二项能够用力的和来取代——直接运用公理 9,即牛顿第二定律。因此,

$$\frac{d\boldsymbol{M}(t)}{dt} = \sum_{i=1}^{N} [\boldsymbol{s}_i(t) - \boldsymbol{b}(t),\ \boldsymbol{F}_i + \boldsymbol{G}_i]$$

但由于 $\boldsymbol{f}_{ij} = -\boldsymbol{f}_{ji}$，取消了内力的力矩,因此,

$$[\boldsymbol{s}_i(t) - \boldsymbol{b}(t),\ \boldsymbol{f}_{ij}] + [\boldsymbol{s}_j(t) - \boldsymbol{b}(t),\ \boldsymbol{f}_{ji}] = [\boldsymbol{s}_i(t) - \boldsymbol{s}_j(t),\ \boldsymbol{f}_{ij}] = 0$$

根据公理 7,证明了这个定理。

守恒定律。如果没有合外力作用于任意质点,即如果对于所有质点 i 和时间 t,$\boldsymbol{G}_i(t)$都是零,那么该质点系统被说成是**封闭的**。对于封闭系统来说,我们有作为定理 3 和 5 直接结果的两个重要的守恒定律：

定理 6. 对于封闭的质点系统,动量 $\boldsymbol{P}(t)$和角动量 $\boldsymbol{M}(t)$守恒,即 $\boldsymbol{P}(t)$和 $\boldsymbol{M}(t)$都是恒定的。

动量守恒定律的某种形式至少追溯到笛卡儿的《哲学原理》(1644),但从对我们有利的观点来看,这种考虑是很大的误解。笛卡儿把运动的量定义为一个物体的大小及其作为标量的速度的乘积。守恒是以这种方式表述的(Ⅱ,36)：①

---

① 这个参考文献是《哲学原理》第二部分的文章 36。所有的引文都取自 V. R. 米勒(V. R. Miller)和 P. R. 米勒(P. R. Miller)的翻译(1983)。

> 上帝是运动的第一原因，他总是在宇宙中保持相等的运动量。

笛卡儿确实想到了大小而不是质量，他在Ⅱ，36、40、43、47－52；Ⅳ，199、203这两篇文章中的陈述支持了这一点。笛卡儿和莱布尼茨之间关于运动量的正确定义的经典争论，实际上直到18世纪才得到解决。像在数学和物理学相关的许多其他问题中，近代的用法接近于欧拉的用法。

328

牛顿以正确的方式非形式地陈述了只受引力作用的物体的动量守恒定律。他在《自然哲学的数学原理》第一卷第Ⅺ节的导言中，是这么说的：

> 如果有更多的物体，一个物体被另一个物体所吸引，再被这些物体所吸引，或者说，它们彼此相互吸引，这些物体在它们中间进行这样的运动，以至于它们共同的重心要么是静止的，要么做匀速直线运动。

角动量守恒是更新近的概念。尽管（例如，在18世纪的刚体理论中）以各种不同的方式运用动量矩或角动量，但是，按照特鲁斯德尔的观点（1968），也许是泊松在1833再版的他的《力学教程》中第一次证明了公理5。

即使不可能对力作出充分的动力学说明，但这些守恒定律在许多问题上都有重要的应用，有时是为了简化独立变量的个数，在其他情况下，是为了提供能用来获得一个系统的大量信息的一般定律。

三体问题中的独立变量的经典简化能够举例说明第一种应用，已经提到过，三体问题被看成是经典力学最著名的问题。在笛卡儿坐标中，我们开始有十八个独立变量——九个位置变量和九个速度变量。动量守恒意味着，该系统的质心做匀速运动，

并且，对于质心的运动来说，这转变为六个常量，因此，变量数简化为十二。角动量守恒另外强加了三个方程，因此，变量数进一步简化为九。最后，下面要讨论的能量守恒致使再增加一个方程，因而整个变量简化为八个。布伦斯(Bruns)表明(1887)，这十个守恒条件是对三体问题的惟一可能的代数约束。注意，总而言之，它们来自经典力学的三个守恒定律。

幸运的是，这些守恒定律推广到了相对论性的(非量子)力学和经典量子力学。一个重要的应用是康普顿效应(Compton effect, 1923)，在这个应用中，没有详细说明整个动力学系统。爱因斯坦早已指出，光子应该有确定的动量，尽管它们的质量为零。康普顿(Compton)认为，对于光子与电子之间的碰撞，动量守恒定律应该成立。因此，光子通过电子的散射应该致使动量转移给电子，从这些光子的散射来看，频率也减小了。康普顿通过实验表明，在X射线的散射实验中，确实发生了频率的减小(康普顿效应)，而且，在做散射实验时能够观察到动量转移给电子。这个事例在方法论意义上的重要性是，没有作出关于相互作用力的定量本性的假设。

在近代之前的物理学中，许多地方隐含了守恒定律的哲学思想。亚里士多德的天体运动理论(在其特征上，天体作为永久的和不变的东西)代表了恒定的或守恒的重要思想。另一个是亚里士多德作为变化基底的质料概念。关于这一点的学术名言是：

329 质料既不能被产生，也不能被破坏，因为所有的产生都产生自质料，所有的毁灭都毁灭为质料。①

---

① 引自Jammer(1961, p. 41)。拉丁语是“Materia non est generabilis nec corruptibilis, quiaomne quod generatur, generatur ex materia, et quod corrumpitur, corrumpitur in materiam.”有些人把这段名言归于托马斯·阿奎那。还有，关于非常类似的陈述，参见亚里士多德的《物理学》第一卷第九章192 a30。

显然,现存的**守恒**这个术语是由莱布尼茨(1692)在他关于**活力**(vis viva)的争论及其固有的定量定义中第一次使用的(在法语中)。[1] 至少,自 19 世纪中叶以来,守恒定律被正确地看成是物理定律中最普遍和最有意义的定律。它们不仅在应用中是强有力的,而且满足鉴别世界上哪个量是不变的和永久的更深层次的哲学需要。动量守恒定律、角动量守恒定律和能量守恒定律也许是最伟大全面的物理学定律,即使对一个孤立的物理系统的详细的动力学结构所知甚少,也能够把这些定律应用于这样的系统。[2]

封闭系统中的嵌入。我现在考虑一个定理,断言质点力学的任何一个系统都能被嵌入到一个封闭系统当中。为了表述我们嵌入的结果,我们需要两个**等价的**力学系统概念。这个等价性概念在某种程度上弱于我们为质点力学系统定义的最严格的同构概念。在我们不要求各个力的结构是一样的而只要求合力相同的意义上,是较弱的。

定义 5. *两个质点力学系统* $\mathfrak{P}=(P, \boldsymbol{s}, m, \boldsymbol{f}, \boldsymbol{g})$ *和* $\mathfrak{P}'=(P', \boldsymbol{s}', m', \boldsymbol{f}', \boldsymbol{g}')$ *是等价的,当且仅当,*

$$P = P'$$

$$\boldsymbol{s} = \boldsymbol{s}'$$

$$m = m'$$

注意,这个等价概念是比“自然的”同构概念既较弱又较强

① 参见莱布尼茨全集,佩尔茨(Pertz)主编,《数学》第六卷,第 217 页。

② 4.1 节的最后与纳脱(1918)的著名定理相联系简要地讨论了守恒定律与不变性之间的关系。能量守恒定律稍后陈述为定理 8,因为守恒系统的条件强于封闭系统的条件,所以必须首先介绍守恒系统的概念。

的几个概念之一；当两个等价系统的各个力没有相同的结构时，是较弱的，但当两个等价系统在运动学的意义上一定是同一的和它们的质量函数也一定是同一的时，是较强的。

我们也需要定义一个质点系统是另一个质点系统的子系统的概念。这个定义采用了针对一个更简单的代数框架内的子系统概念已经给出的讨论。

定义 6. 设 $\mathfrak{P}=(P, \mathbf{s}, m, \mathbf{f}, \mathbf{g})$ 是一个质点力学系统，设 P' 是 P 的一个非空子集，设 $\mathbf{s}'$ 和 m' 是前面的自变量限于 P' 的 330 $\mathbf{s}$ 和 m 的函数，设 $\mathbf{f}'$ 是前两个自变量限于 P' 的 $\mathbf{f}$ 的函数，再设 $\mathbf{g}'$ 是函数，使得对于 P' 中的每个 p 和 T 中的 t，

$$\mathbf{g}'(p, t, 1) = \sum_{q \in P-P'} \mathbf{f}(p, q, t)$$

并且对于每个 i

$$\mathbf{g}'(p, t, i+1) = \mathbf{g}(p, t, i)$$

于是，$\mathfrak{P}'=(P', \mathbf{s}', m', \mathbf{f}', \mathbf{g}')$ 是 $\mathfrak{P}$ 的一个子系统。

这个嵌入表征定理恰好就是下列简单结果。

定理 7. 每个质点力学系统都等价于一个封闭的质点力学系统的子系统。

这个定理的证明可以在苏佩斯的著作(Suppes, 1957/1999: p.302 - 304)中找到。因此，这里省略了。从物理学的观点来看，这种解释带来了一个相当有趣的问题。这类嵌入定理在数学上是标准的，但它们在物理学中的意义更有限，理由很简单，我们嵌入一个已知系统的更大的封闭系统，不是由真正的物理质点而是由想象的或纯粹的概念化的质点所增大的初始系统构成的。换言之，当我们把一个系统嵌入到一个封闭系统中时，我们不保证，在更大的封闭系统中，附加的质点是实际存在的。然而，这个定理的概念点与数学中的一般情况依然相同。我们

已经表明,在受任意外力作用的任意的质点力学系统中可能产生的任何类型的运动,都能够在一个扩大了的封闭的质点力学系统中得到复制和表征。因此,为了说明宇宙中的具体物体的所有运动的起源,从关于力学的适当性的概念论证的角度来看,这个定理是有意义的。这个定理所表明的是,根据内力满足力的大小相等方向相反的牛顿第三定律,我们通过概念化,总能解释一个系统的运动。

守恒系统。至少在某种程度上指出,最重要的是,用势能函数表征一个系统的质点受到的作用力。这个概念是由拉格朗日于 1773 年提出的(Oeuvres, Ⅵ, p. 335),势的命名归功于英国数学家和科学家乔治·格林(George Green)。

定义 7. *设 $\mathfrak{P}=(P, T, m, s, \boldsymbol{f}, \boldsymbol{g})$ 是有 n 个质点的质点力学系统。于是,这个系统是守恒的,当且仅当,存在着一个从 E^{3n} 到 R 的可微函数 U($\mathfrak{P}$ 的势能函数),使得对于 $1 \leqslant i \leqslant n$ 和 T 中的 t,*

$$\boldsymbol{F}_i(t)=\frac{\partial U}{\partial \boldsymbol{q}_i}(\boldsymbol{q}_1(t), \cdots, \boldsymbol{q}_n(t)) \quad i=1, \cdots, n$$

正如这个定义所表明的,可用势能函数的偏导数来表示一个守恒系统中的合力。以这种形式给出这个定义强调了这种观点,但传统表述是,把 $m_i \ddot{\boldsymbol{q}}_i$ 放在这个定义方程的左边,而不是 $F_i(t)$。在当前的框架中,当我们从力和牛顿第二定律开始时,这两种表述显然是等价的。 331

下面是一些简单的但重要的守恒系统的例子。

落石。这个问题是一维的,用 $x(t)$ 表示石头在时间 t 离开地面的高度。运动方程是

$$m\ddot{x}=-gm$$

其中,m 是石头的质量,g 是引力常数,同时,g 近似等于 9.8 米/秒。于是

$$U(x)=gmx$$

谐振子。由一个质点组成的这个简单的力学系统在经典力学和量子力学中都有许多不同的应用。经典的例子有:小振动的单摆,在弹簧秤中所用的这类弹簧的小延伸。运动的一维方程是

$$m\ddot{x}=-m\alpha^2 x$$

其中,m 是质点的质量,α 是物理常数,从一种应用到另一种应用这个常数不同。这里,

$$U(x)=\frac{m\alpha^2 x^2}{2}$$

而且,我们能把运动方程写为

$$m\ddot{x}=\frac{\partial U(x)}{\partial x}$$

三体问题。假如三个质点的质量是 m_1、m_2 和 m_3,它们仅有的力是彼此间的引力,关于这个著名的三体问题,有下列势能函数:

$$U(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3)=-\frac{m_1 m_2}{|\boldsymbol{q}_1-\boldsymbol{q}_2|}-\frac{m_2 m_3}{|\boldsymbol{q}_2-\boldsymbol{q}_3|}-\frac{m_3 m_1}{|\boldsymbol{q}_3-\boldsymbol{q}_1|}$$

我们现在希望证明的一个重要定理是,这个守恒系统的总能量是守恒的。首先,对于由 n 个质点组成的任意系统,作为时间函数的动能 $T(t)$ 定义为

$$T(t)=\frac{1}{2}\sum_{i=1}^{n} m_i \dot{\boldsymbol{q}}_i^2$$

这个量在力学中有长久而复杂的历史。从本质上看,它首先是由莱布尼茨提出的(*Acta erud.*, 1695)。他称一个质点的质量乘以它的速度的平方为 *vis viva*(活力)。具有势能函数 U 的一个守恒系统的总能量 E 是

$$E = T + U$$

其中,省略了函数的自变量。更明确地说,

$$E(t) = T(t) + U(\boldsymbol{q}_i(t), \cdots, \boldsymbol{q}_n(t))$$

定理 8. (能量守恒定律)一个守恒系统的总能量不随时间的变化而变化。 332

证明: 运用上面给出的 T 和 U 的定义,不难证明

$$\frac{d}{dt}(T + U) = 0$$

下面的定理与封闭系统——无外力——和守恒系统有关。证明省略了。

定理 9. 如果一个质点系统是封闭的,而且,相互作用的内力只取决于质点间的距离,即对于所有质点 i 与 j 和时间 t,

$$\boldsymbol{f}(i, j, t) = \boldsymbol{f}(|\boldsymbol{q}_i - \boldsymbol{q}_j|)$$

于是,这个系统是守恒的,即它有一个势能函数。

7.2 量子力学中的隐变量的表征定理

现在,关于量子力学中的隐变量的文献非常之多。即使在隐变量可能是决定论的变量时,本节也主要研究隐变量的概率表征定理的部分文献。幸运的是,没有量子力学的大量知识,也

能理解这些主要结果,这里不从头开始阐述。①

首先,我们陈述和略述我们考虑的集合的基本定理的证明:对于一个有穷的或连续的可观察量的集合,即用概率论的语言来说,对于一个随机变量的集合,存在着一个因子隐变量(factoring hidden variable),即在形式上的一种共同原因(common cause),当且仅当,这些可观察量有一个联合概率分布。这个定理在物理上的重要方面是,在很一般的条件下,存在的隐变量能够完全被简化为只是可观察量之间的相互关系,也就是说,确定它们是否有与给出的数据(例如,可观察量的平均值、方差和关联)相一致的联合概率分布的问题。

我们强调指出,尽管大多数文献只限于二阶矩,比如,协方差和关联,但是,根本没有必要作出这样一种限制。事实上,在著名的格林伯格(Greenberger)、霍恩(Horne)和蔡林格(Zeilinger)(1989)根据三个和四个质点位形提供的关于隐变量的新的思想实验(Gedanken experiment)产生的三阶或四阶矩违反了这种限制。我们关于抽象的GHZ结果的概率证明,参见定理7,关于相关的不等式的证明,参见定理8。

众所周知,贝尔(Bell)关于隐变量的结果在大多数情况下限于±1的可观察量,比如,自旋或极化。但这种限制是不必要的。我们的一般结果涵盖了任何有穷的或连续的可观察量。最333 后,我们给出了关于高斯可观察量的隐变量的不同结果,并把下列非线性不等式表述为最终定理,这一不等式对三个高斯随机变量拥有与它们已知的平均值、方差和关联相融的联合分布来

① 已有的许多结果来自阿卡西奥·德·巴罗斯(Acacio de Barros)和加里·奥斯(Gary Oas)的共同工作(Suppes, de Barros and Oas 1998; de Barros and Suppes 2000)。

说既是必要的，也是充分的。

因式分解。在关于隐变量的文献中，因式分解原理（the principle of factorization）有时被取名为定域性原理（principle of locality）。这个术语实际上是无关紧要的，但意义很关键。我们想到了关于（连续的或离散的）随机变量的一个相当一般的原理，这个原理是：设 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 是随机变量，于是，存在着一个随机变量 $\boldsymbol{\lambda}$（它是所预期的隐变量，使得 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 在条件上与已知的 $\boldsymbol{\lambda}$ 无关）的充分必要条件是，在不考虑 $\boldsymbol{\lambda}$ 的前提下，存在着 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 的一个联合概率分布。这是我们的第一个定理，它是与隐变量和可观察的随机变量的联合概率分布相关的一般的基本定理。

定理 1.（Suppes and Zanotti 1981；Holland and Rosenbaum 1986）。*设已知 n 个有穷的或连续的随机变量 $X_1, \cdots, X_n$。于是，存在着一个隐变量 $\boldsymbol{\lambda}$，使得存在着具有特性（$\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n, \boldsymbol{\lambda}$）的一个联合概率分布 F：*

(i) $F(x_1, \cdots, X_n \mid \lambda) = P(\boldsymbol{X}_1 \leqslant x_1, \cdots, \boldsymbol{X}_n \leqslant x_n \mid \boldsymbol{\lambda} = \lambda)$

(ii) *条件无关成立，即对于所有的 $x_1, \cdots, x_n, \lambda$ 来说，*

$$F(x_1, \cdots, x_n \mid \lambda) = \prod_{j=1}^{n} F_j(x_j \mid \lambda)$$

当且仅当，存在 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 的一个联合概率分布。此外，为了成为决定论的，可以构造 $\boldsymbol{\lambda}$，即，已知 $\boldsymbol{\lambda}$，每个 $\boldsymbol{X}_i$ 的条件方差都是零。

为了完全明确起见用符号表示为

$$F_j(x_j \mid \lambda) = P(\boldsymbol{X}_j \leqslant x_j \mid \boldsymbol{\lambda} = \lambda) \tag{1}$$

证明思路。考虑三个 ± 1 的随机变量 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$。有八种可能的联合结果（$\pm 1, \pm 1, \pm 1$）。设 p_{ijk} 是结果（i, j, k）的概率。把这个概率赋予我们构造的隐变量 $\boldsymbol{\lambda}$ 的值 λ_{ijk}。于是，四

元组(i, j, k, λ_{ijk})的概率恰好是 p_{ijk}，而且，条件概率是决定论的，即

$$P(\boldsymbol{X}=i, \boldsymbol{Y}=j, \boldsymbol{Z}=k \mid \lambda_{ijk})=1$$

并且因式分解是直接的，即

$$\begin{aligned} & P(\boldsymbol{X}=i, \boldsymbol{Y}=j, \boldsymbol{Z}=k \mid \lambda_{ijk}) \\ = & P(\boldsymbol{X}=i \mid \lambda_{ijk}) P(\boldsymbol{Y}=j \mid \lambda_{ijk}) P(\boldsymbol{Z}=k \mid \lambda_{ijk}) \end{aligned}$$

把这条论证的思路推广到一般情况证明了，可观察量的联合概率分布对于存在的因子隐变量，即共同的原因来说是充分的。从定理 1 的表述来看，必要性是明显的，因为$(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n)$的联合分布是更大分布$(\boldsymbol{X}_1 \cdots, \boldsymbol{X}_n, \boldsymbol{\lambda})$的边际分布。

显然，$\boldsymbol{\lambda}$ 的构造纯粹是数学的。它本身没有物理内容。事334 实上，证明本身很简单。在为了拥有一种联合概率分布给出可观察变量的有效标准时，所有真正的数学困难都将会出现。正如我们后面要详细评论的那样，我们还没有很好的标准来判断，在这种形式的不等式中，有 $m>2$ 的有穷值的三个随机变量的一种联合分布的必要条件和可能的充分条件，像在更高自旋的情况中那样。

当我们把额外的物理假设强加于隐变量 $\boldsymbol{\lambda}$ 时，$\boldsymbol{\lambda}$ 的物理内容就超越了这些可观察量的联合分布的界限。在关于两个随机变量的下列定理中包含了一个简单的例子。我们把额外的对称性条件强加给条件期望值，然后，只有当这两个可观察量的联合分布不是负的时(关于联合分布的一个强的附加限制)，才存在着一个隐变量。这个表征定理的证明能在引证其陈述的文章中找到。

定理 2. (Suppes and Zanotti 1980)。**设 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 是有两个值的随机变量，为了明确起见，可能值是 1 与 −1，并有正方差，**

即 $\sigma(\boldsymbol{X})$、$\sigma(\boldsymbol{Y}) > 0$。此外，设 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 是可互换的，即

$$P(\boldsymbol{X}=1, \boldsymbol{Y}=-1) = P(\boldsymbol{X}=-1, \boldsymbol{Y}=1)$$

于是，存在一个隐变量 $\boldsymbol{\lambda}$，使得

$$E(\boldsymbol{XY} \mid \boldsymbol{\lambda}=\lambda) = E(\boldsymbol{X} \mid \boldsymbol{\lambda}=\lambda)E(\boldsymbol{Y} \mid \boldsymbol{\lambda}=\lambda)$$

和

$$E(\boldsymbol{X} \mid \boldsymbol{\lambda}=\lambda) = E(\boldsymbol{Y} \mid \boldsymbol{\lambda}=\lambda)$$

对于每个值 λ（零测度集合上的可能值除外），这个充分必要条件是，$\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的关联是非负的。

通常，在物理学中，像在本节中一样，我们惟一感兴趣的是平均值、方差和协方差——所谓的二阶概率理论，因为我们只考虑二阶矩。我们说，相对于存在头两个矩的随机变量 $\boldsymbol{X}_1$，…，$\boldsymbol{X}_n$，一个隐变量 $\boldsymbol{\lambda}$ 满足**二阶因式分解条件**，当且仅当，

$$\text{(a) } E(\boldsymbol{X}_1 \cdots \boldsymbol{X}_n \mid \boldsymbol{\lambda}) = E(\boldsymbol{X}_1 \mid \boldsymbol{\lambda}) \cdots E(\boldsymbol{X}_n \mid \boldsymbol{\lambda})$$

$$\text{(b) } E(\boldsymbol{X}_1^2 \cdots \boldsymbol{X}_n^2 \mid \boldsymbol{\lambda}) = E(\boldsymbol{X}_1^2 \mid \boldsymbol{\lambda}) \cdots E(\boldsymbol{X}_n^2 \mid \boldsymbol{\lambda})$$

我们就有作为定理 1 的直接推论的下列表征。

定理 3. 设已知 n 个离散的或连续的随机变量。如果存在着 $\boldsymbol{X}_1$，…，$\boldsymbol{X}_n$ 的一个联合概率分布，那么存在着一个决定论的隐变量 $\boldsymbol{\lambda}$，使得 $\boldsymbol{\lambda}$ 满足相对于 $\boldsymbol{X}_1$，…，$\boldsymbol{X}_n$ 的二阶因式分解条件。

我们称为**因式分解定理**的定理 1 和定理 3 的非形式陈述是，存在一个因子隐变量 $\boldsymbol{\lambda}$ 的充分必要条件，恰好是存在已知随机变量 $\boldsymbol{X}_i$ 的一个联合概率分布。在我们的观点中，对于隐变量来说，因式分解条件通常是太强的条件。一个惊人的例子是，经典力学中的引力现象满足定域性条件，但不满足因式分解条件。

定域性。我们下面想要讨论的系统概念是定域性。我们说 335

的定域性的意思是下面引自约翰·贝尔 1964 年的著名文章(Bell,1964)中所指的定域性的意思。

> 正是定域性要求,更精确地说,曾经相互作用过的两个系统,分开之后,对一个系统的测量结果不受远距离的另一个系统的操作的影响,引起了本质上的困难……关键的假设是,粒子 2 的结果 B 不依赖于设置 $\boldsymbol{a}$,即粒子 1 的磁场,A 也不依赖于 $\boldsymbol{b}$。

尽管他们在一个抽象的层次上陈述定理 1 和定理 3,没有参照任何时空或其他物理学的考虑,但是,在他们的陈述中,存在着一个隐含的定域性假设。为了使定域性假设更明确,我们需要用到附加概念。对于每个随机变量 $\boldsymbol{X}_i$ 来说,我们为用来测量随机变量 $\boldsymbol{X}_i$ 的值的局部仪器(在时空中)引入一个参数向量 M_i。

定义 1.(定域性条件Ⅰ)

$$E(\boldsymbol{X}_i^k \mid M_i, M_j, \boldsymbol{\lambda}) = E(\boldsymbol{X}_i^k \mid M_i, \boldsymbol{\lambda})$$

其中,$k = 1, 2$,对应于 $\boldsymbol{X}_i$ 的前两个矩,$i \neq j$ 且 $1 \leqslant i, j \leqslant n$。注意,我们只考虑 M_j,假定在已知的实验操作中,只研究 $\boldsymbol{X}_i$ 和 $\boldsymbol{X}_j$ 的关系。显然可推广到更多的变量。在许多实验中,测量仪器的方向是最重要的参数,它是 M_i 的一个分量。

定义 2.(定域性条件Ⅱ)**$\boldsymbol{\lambda}$ 的分布独立于 M_i 和 M_j 的参数值,即对于所有函数 g,期望值 $E(g(\boldsymbol{\lambda}))$ 和 $E(g(\boldsymbol{\lambda}) \mid M_i, M_j)$ 都是有穷的,**

$$E(g(\boldsymbol{\lambda})) = E(g(\boldsymbol{\lambda}) \mid M_i, M_j)$$

这里我们承袭苏佩斯的文章(1976)。根据定理 3,要求在条件Ⅰ意义上的定域性才能满足这样的假设:每个 $\boldsymbol{X}_i$ 都有固定的平均值和方差。如果当 $\boldsymbol{X}_i$ 与 $\boldsymbol{X}_j$ 组成一对时,对 $\boldsymbol{X}_i$ 的实验观

察，不同于当 $\boldsymbol{X}_i$ 与 $\boldsymbol{X}_{j'}$ 组成一对时，对 $\boldsymbol{X}_i$ 的观察结果，那么就违反了平均值和方差的不变性假设。在构造 $\boldsymbol{\lambda}$ 时，必须满足定域性条件Ⅱ的限制，而且，很容易核实，情况确实如此。

我们在定理 4 中包含了这些评论。

定理 4. 设已知 n 个随机变量 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ 满足定理 3 的假设。设 M_i 是测量 $\boldsymbol{X}_i$ 时的定域的参数向量，再设每个 $\boldsymbol{X}_i$ 都满足定域性条件Ⅰ。于是，如果存在一个联合概率分布 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$，就存在一个隐变量 $\boldsymbol{\lambda}$ 满足定域性条件Ⅱ和二阶因式分解条件。

下一个定理陈述了两个条件，这两个条件等价于苏佩斯和扎诺蒂(1981)为只有两个值的三个随机变量提供的一个不等式条件。

定理 5. 设已知具有值±1 的三个随机变量 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 满足对称性条件 $E(\boldsymbol{X}) = E(\boldsymbol{Y}) = E(\boldsymbol{Z}) = 0$，并且具有已知的协方差 $E(\boldsymbol{XY})$、$E(\boldsymbol{YZ})$和 $E(\boldsymbol{XZ})$。于是，下列三个条件是等价的：

(i) 存在着满足定域性条件Ⅱ的一个隐变量 $\boldsymbol{\lambda}$，而且，二阶因式分解条件的方程(a)成立。 336

(ii) 存在着与已知的平均值和协方差相一致的随机变量 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 的一种联合概率分布。

(iii) 随机变量 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 满足下列方程：

$$\begin{aligned}-1 &\leqslant E(\boldsymbol{XY}) + E(\boldsymbol{YZ}) + E(\boldsymbol{XZ}) \\ &\leqslant 1 + 2\mathrm{Min}(E(\boldsymbol{XY}), E(\boldsymbol{YZ}), E(\boldsymbol{XZ}))\end{aligned}$$

关于这个定理，特别是，在(iii)中给出的不等式，有几种评论。第一种观点是，这些不等式如何与下列著名的贝尔不等式(1964)联系起来：

$$1 + E(\boldsymbol{YZ}) \geqslant |E(\boldsymbol{XY}) - E(\boldsymbol{XZ})| \quad (2)$$

对于具有值±1与期望值等于零的随机变量 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 的一个联合概率分布的存在性来说，贝尔不等式事实上既是不必要的，也是不充分的。不充分性很容易从使三个协方差都等于 $-\frac{1}{2}$ 时看出。于是，贝尔不等式得到了满足，因为

$$1-\frac{1}{2} \geqslant \left|-\frac{1}{2}-\left(-\frac{1}{2}\right)\right|$$

即

$$\frac{1}{2} \geqslant 0$$

但是，正如从(iii)清楚地看到的那样，根本不存在三个协方差都等于 $-\frac{1}{2}$ 的联合概率分布，因为

$$-\frac{1}{2}+-\frac{1}{2}+-\frac{1}{2}<-1$$

第二，贝尔不等式是不必要的。设 $E(\boldsymbol{XY})=\frac{1}{2}$，$E(\boldsymbol{XZ})=-\frac{1}{2}$ 和 $E(\boldsymbol{YZ})=-\frac{1}{2}$，于是，违反(2)，因为

$$1-\frac{1}{2}<\left|\frac{1}{2}-\left(-\frac{1}{2}\right)\right|$$

但满足(iii)，因此，存在一个联合分布

$$-1 \leqslant \frac{1}{2}-\frac{1}{2}-\frac{1}{2} \leqslant 1+2 \operatorname{Min}\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)$$

即

$$-1 \leqslant-\frac{1}{2} \leqslant 0$$

贝尔通过一个定域隐变量理论所满足的某些情形，推导出他的不等式，但由于量子力学的协方差等于 $-\cos\theta_{ij}$，违反了这个不等式。特别是，设 $\theta_{\boldsymbol{XY}} = 30°$，$\theta_{\boldsymbol{XZ}} = 60°$，$\theta_{\boldsymbol{YZ}} = 30°$，因此，在几何意义上，$\boldsymbol{Y}$ 平分 $\boldsymbol{X}$ 和 $\boldsymbol{Z}$。于是

$$\left|-\frac{1}{2}-\left(-\frac{\sqrt{3}}{2}\right)\right|>1-\frac{\sqrt{3}}{2}$$

第二个评论是，对于具有期望值等于零的三值随机变量来说，(iii)是不必要的。设三个值是 1，0，-1。这里是一个反例，其中，三种关联中的每一种关联都是 $-\frac{1}{2}$，因此，根据和等于 $-\frac{3}{2}$，违反了(iii)。注意，对于具有期望值是零的 ± 1 的变量来说，协方差和关联是相等的，因为方差是 1。一般情况下，不是这种情形。而且，尤其是，我们的反例并不是这种情形。除了所提到的特殊情况之外，为了具有严格的通用性，应该根据关联而不是协方差来写不等式(例如，参见下面的定理 10)。 337

存在着具有下列值的一种联合概率分布。设 $p(x, y, z)$ 是三个值都已知(例如 1，-1，0)的概率。于是，对于所有的 x、y 和 z，我们当然有

$$p(x, y, z) \geqslant 0 \text{ 和 } \sum_{x, y, z} p(x, y, z) = 1$$

其中，x、y 和 z 每一个都有三个值 1，0，-1。因此，设

$$\begin{aligned} p(-1, 0, 1) &= p(1, -1, 0) = p(0, 1, -1) \\ &= p(1, 0, -1) = p(-1, 1, 0) \\ &= p(0, -1, 1) = \frac{1}{6} \end{aligned}$$

而且，其他 21 个概率 $p(x, y, z) = 0$。于是，不难证明，在这个

模型中，$E(\boldsymbol{X})=E(\boldsymbol{Y})=E(\boldsymbol{Z})=0$，$\mathrm{Var}(\boldsymbol{X})=\mathrm{Var}(\boldsymbol{Y})=\mathrm{Var}(\boldsymbol{Z})=\frac{2}{3}$，而且，$\mathrm{Cov}(\boldsymbol{XY})=\mathrm{Cov}(\boldsymbol{YZ})=\mathrm{Cov}(\boldsymbol{XZ})=-\frac{1}{3}$，因此，关联是

$$\rho(\boldsymbol{XY})=\rho(\boldsymbol{YZ})=\rho(\boldsymbol{XZ})=-\frac{1}{2}$$

这三种关联之和就是$-\frac{3}{2}$。

有点令人沮丧的数学事实是，即使是对于具有 n 个值和期望值等于零的三个随机变量来说，为了拥有与已知的平均值、方差和协方差相一致的一个联合概率分布，一种不同的研究也似乎需要为每个 n 寻找充分必要条件。一个更一般的递归结果是很令人向往的，但似乎是未知的。这样的结果与多值自旋现象的研究相关。

下一个著名的定义陈述等价于只有两个值的随机变量的贝尔不等式的两个条件。这种形式的不等式归功于克劳泽(Clauser)、霍恩、希莫尼(Shimony)和霍尔特(Holt)的工作(1969)，称之为 CHSH。(ii)和(iii)的等价归功于法恩(Fine)的工作(1982)。

定理 6. (贝尔不等式)。设已知 n 个随机变量满足定理 4 的定域性假设。设 $n=4$，即随机变量的个数，设每个 $\boldsymbol{X}_i$ 都是离散的，其值为 ± 1，设满足对称性假设 $E(\boldsymbol{X}_i)=0(i=1,\cdots,4)$，设 $\boldsymbol{X}_1=\boldsymbol{A},\boldsymbol{X}_3=\boldsymbol{B},\boldsymbol{X}_2=\boldsymbol{A}',\boldsymbol{X}_4=\boldsymbol{B}'$，已知具有协方差 $E(\boldsymbol{AB})$、$E(\boldsymbol{AB}')$ $E(\boldsymbol{A}'\boldsymbol{B})$和 $E(\boldsymbol{A}'\boldsymbol{B}')$。于是，下列三个条件是等价的。

(i) 存在着满足定域性条件 Ⅱ 的一个隐变量 $\boldsymbol{\lambda}$，而且，二阶因式分解条件的方程(a)成立。

(ii) 存在着与已知的平均值和协方差相一致的随机变量 **A**、**A**′、**B** 和 **B**′的一个联合概率分布。

(iii) 随机变量 **A**、**A**′、**B** 和 **B**′满足 CHSH 形式中的贝尔不等式：

$$-2 \leqslant E(\boldsymbol{AB}) + E(\boldsymbol{AB}') + E(\boldsymbol{A'B}) - E(\boldsymbol{A'B'}) \leqslant 2$$

$$-2 \leqslant E(\boldsymbol{AB}) + E(\boldsymbol{AB}') - E(\boldsymbol{A'B}) + E(\boldsymbol{A'B'}) \leqslant 2$$

$$-2 \leqslant E(\boldsymbol{AB}) - E(\boldsymbol{AB}') + E(\boldsymbol{A'B}) + E(\boldsymbol{A'B'}) \leqslant 2$$

$$-2 \leqslant - E(\boldsymbol{AB}) + E(\boldsymbol{AB}') + E(\boldsymbol{A'B}) + E(\boldsymbol{A'B'}) \leqslant 2$$

现在，我们将表明，对于有三个值的随机变量（自旋为 1 的粒子）来说，CHSH 的不等式仍然是有效的。考虑具有 3 -态可观察量的一个自旋为 1 的粒子，$A(a, \lambda) = +1, 0, -1$，$B(b, \lambda) = +1, 0, -1$。$\lambda$ 是具有归一化概率密度 $\rho(\lambda)$的一个隐变量。这些可观察量的期望值定义为 338

$$E(a, b) = \int AB\rho(\lambda)d\lambda \tag{3}$$

为了明确起见，我们禁止这个变量依赖于 A 和 B。考虑下列差

$$|E(a, b) - E(a, b')| = \left|\int A[B - B']\rho(\lambda)d\lambda\right| \tag{4}$$

既然密度 $\rho > 0$ 和 $|A| = 1, 0$，所以，我们有下列不等式

$$|E(a, b) - E(a, b')| \leqslant \int |A[B - B']| \rho(\lambda)d\lambda \tag{5}$$

$$\leqslant \int |[B - B']| \rho(\lambda)d\lambda \tag{6}$$

同样，我们有下列不等式

$$|E(a', b) + E(a', b')| = \left|\int A'[B + B']\rho(\lambda)d\lambda\right| \tag{7}$$

$$\leqslant \int |[B + B']| \rho(\lambda) d\lambda \quad (8)$$

增加两个表达式,我们获得下列不等式

$$|E(a, b) - E(a, b')| + |E(a', b) + E(a', b')|$$
$$= \int (|[B - B']| + |[B + B']|) \rho(\lambda) d\lambda \quad (9)$$

除了当 B 和 B' 都等于零之外,在所有情况下,方括号中的项都等于 2;在那里,右边的项化为零。根据这一点和隐变量密度的归一化条件,我们有与自旋-$\frac{1}{2}$的 CHSH 不等式相同的不等式

$$|E(a, b) - E(a, b')| + |E(a', b) + E(a', b')| \leqslant 2 \quad (10)$$

注意,我们能够通过把函数 $2(|E(a, b)| - 1)(|E(a, b')| - 1)$ 加到左边,产生一个更强的不等式。

对于更高的自旋来说,我们类似地进行下去,就推导出自旋-j粒子必须满足的下列不等式

$$|E(a, b) - E(a, b')| + |E(a', b) + E(a', b')| \leqslant 2j \quad (11)$$

如果我们定义归一化可观察量$\frac{A(a, \lambda)}{j}$,定域隐变量理论就需要满足原始的 CHSH 不等式,尽管能够构造更强的不等式。

在佩雷斯(Peres)关于更高自旋粒子的工作中,从多态算符映射到一个两态算符,来定义可观察量(Peres, 1992)。在这种映射条件下,已经表明,检测器的特定参数的设置违反了贝尔不等式。

GHZ 型实验。我们改变关注点,现在首先考虑 GHZ 型的三粒子实验。我们知道所有的论证,特别是 GHZ(Greenberger, Horne and Zeilinger 1989)自己的论证,即在格林伯格、霍恩、希莫尼和蔡林格(1990)以及梅尔曼(Mermin,1990a,1993)的工作中通过下列方式进行的更加推广了的论证:即假设存在着一个决定论的隐变量,然后,推导出一种矛盾。从定理 1 直接得出,隐变量的不存在性等价于已知可观察的随机变量的一个联合概率分布的不存在性。下一个定理陈述这种纯概率的 GHZ 结果,更重要的是,这个证明完全是根据可观察量进行的,没有考虑可能的隐变量。 339

定理 7. 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 和 $\boldsymbol{C}$ 是具有一个联合概率分布的 ± 1 的随机变量,使得 $E(\boldsymbol{A}) = E(\boldsymbol{B}) = E(\boldsymbol{C}) = 1$。于是,$E(\boldsymbol{ABC}) = 1$。

证明: 既然 $E(A) = 1$,所以,

$P(\bar{a}) = P(\bar{a}bc) = P(\bar{a}b\bar{c}) = P(\bar{a}\bar{b}c) = P(\overline{abc}) = 0$,其中,$P(\bar{a}bc) = P(\boldsymbol{A} = -1, \boldsymbol{B} = 1, \boldsymbol{C} = 1)$ 等等,根据 $E(\boldsymbol{B})$和 $E(\boldsymbol{C})$的类似论证,我们有 $P(abc) = 1$,它直接蕴含了所希望的结果。

我现在采纳梅尔曼(1990b)给出的量子力学论证。我们从对 GHZ 型实验的三粒子纠缠态的描述开始。

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1\ |+\rangle_2\ |+\rangle_3 + |-\rangle_1\ |-\rangle_2\ |-\rangle_3) \quad (12)$$

这个态是下列自旋算符的一个本征态:

$$\boldsymbol{A} = \hat{\sigma}_{1x}\hat{\sigma}_{2y}\hat{\sigma}_{3y},\ \boldsymbol{B} = \hat{\sigma}_{1y}\hat{\sigma}_{2x}\hat{\sigma}_{3y} \quad (13)$$

$$\boldsymbol{C} = \hat{\sigma}_{1y}\hat{\sigma}_{2y}\hat{\sigma}_{3x},\ \boldsymbol{D} = \hat{\sigma}_{1x}\hat{\sigma}_{2x}\hat{\sigma}_{3x} \quad (14)$$

如果我们根据量子力学计算上述关联的期望值,我们立即得到,$E_Q(\hat{\boldsymbol{A}}) = E_Q(\hat{\boldsymbol{B}}) = E_Q(\hat{\boldsymbol{C}}) = 1$ 和 $E_Q(\hat{\boldsymbol{D}}) = -1$。(这个过程显

示的所有细节太长，无法包括，但在量子力学的语境中，这个论证是基本的和重要的。)此外，

$$E_Q(\boldsymbol{ABC}) = (s_{1x}s_{2y}s_{3y})(s_{1y}s_{2x}s_{3y})(s_{1y}s_{2y}s_{3x}) \tag{15}$$

$$= s_{1x}s_{2x}s_{3x}(s_{1y}^2 s_{2y}^2 s_{3y}^2) \tag{16}$$

既然 s_{ij} 只能是 1 或 -1，所以，$s_{1y}^2 = s_{2y}^2 = s_{3y}^2 = 1$，而且，我们从(16)得出

$$E_Q(\boldsymbol{ABC}) = s_{1x}s_{2x}s_{3x} = E_Q(\boldsymbol{D}) \tag{17}$$

因此，我们有了一个矛盾。在经典意义上，

$$E(\boldsymbol{ABC}) = 1$$

但在量子力学的意义上，

$$E_Q(\boldsymbol{ABC}) = -1$$

从上面的推导来看，显然，如果我们允许 $\boldsymbol{\lambda}$ 依赖于实验设置，即，如果我们允许 $\boldsymbol{\lambda}$ 是一个语境的隐变量，人们就能避免矛盾。换言之，GHZ 定理所证明的是，非语境的隐变量不可能重复量子力学的预言。

检测器效率低。然而，GHZ 预言的这个惊人特征有一个主要问题。既然在实验的意义上人们不可能获得完全关联的事件，人们如何能基于关联陈述从实验上证实预言呢？幸运的是，正如我们在下列定理及其推论中所表明的那样，在 GHZ 态出现的关联是如此强，以至于即使我们允许有实验误差，特别不可避免的是检测器效率低，仍然能证实不存在一种联合分布。

340

定理 8. (de Barros and Suppes 2000)。*如果 $\boldsymbol{A}$、$\boldsymbol{B}$ 和 $\boldsymbol{C}$ 是三个 ± 1 的随机变量，对于已知的期望值 $E(\boldsymbol{A})$、$E(\boldsymbol{B})$、$E(\boldsymbol{C})$ 和 $E(\boldsymbol{ABC})$ 来说，存在着一个联合概率分布，当且仅当，满足下列不等式：*

$$-2 \leqslant E(\mathbf{A}) + E(\mathbf{B}) + E(\mathbf{C}) - E(\mathbf{ABC}) \leqslant 2 \qquad (18)$$

$$-2 \leqslant E(\mathbf{A}) + E(\mathbf{B}) - E(\mathbf{C}) + E(\mathbf{ABC}) \leqslant 2 \qquad (19)$$

$$-2 \leqslant E(\mathbf{A}) - E(\mathbf{B}) + E(\mathbf{C}) + E(\mathbf{ABC}) \leqslant 2 \qquad (20)$$

$$-2 \leqslant -E(\mathbf{A}) + E(\mathbf{B}) + E(\mathbf{C}) + E(\mathbf{ABC}) \leqslant 2 \qquad (21)$$

证明：首先，我们证明必要性。让我们假设，有由八个原子事件 abc、$ab\bar{c}$、$a\bar{b}c$，…，$\bar{a}\bar{b}\bar{c}$ 组成的一个联合概率分布。于是，

$$E(\mathbf{A}) = P(a) - P(\bar{a})$$

其中，

$$P(a) = P(abc) + P(a\bar{b}c) + P(ab\bar{c}) + P(a\bar{b}\bar{c})$$

和

$$P(\bar{a}) = P(\bar{a}bc) + P(\bar{a}\bar{b}c) + P(\bar{a}b\bar{c}) + P(\bar{a}\bar{b}\bar{c})$$

同样的方程对于 $E(\mathbf{B})$ 和 $E(\mathbf{C})$ 也成立。对于 $E(\mathbf{ABC})$，我们得到

$$\begin{aligned} E(\mathbf{ABC}) &= P(\mathbf{ABC} = 1) - P(\mathbf{ABC} = -1) \\ &= P(abc) + P(a\bar{b}\bar{c}) + P(\bar{a}\bar{b}c) + P(\bar{a}b\bar{c}) \\ &\quad - [P(a\bar{b}c) + P(ab\bar{c}) + P(\bar{a}bc) + P(\bar{a}\bar{b}\bar{c})] \end{aligned}$$

与上面的第一个不等式相对应，我们现在总结期望值的概率表达式

$$F = E(\mathbf{A}) + E(\mathbf{B}) + E(\mathbf{C}) - E(\mathbf{ABC})$$

而且，获得表达式

$$\begin{aligned} F = &\, 2[P(abc) + P(\bar{a}bc) + P(a\bar{b}c) + P(ab\bar{c})] \\ &- 2[P(\bar{a}\bar{b}\bar{c}) + P(\bar{a}\bar{b}c) + P(\bar{a}b\bar{c}) + P(a\bar{b}\bar{c})] \end{aligned}$$

既然所有的概率都是非负的，而且，和小于等于 1，我们立即推

出不等式(18)。其他三个不等式的推导很类似。

证明逆命题,即这些不等式意味着一个联合概率分布的存在性,稍微更复杂些。我们限于对称情况。

$$P(a) = P(b) = P(c) = p$$

$$P(\boldsymbol{ABC} = 1) = q$$

因此,

$$E(\boldsymbol{A}) = E(\boldsymbol{B}) = E(\boldsymbol{C}) = 2p - 1$$

$$E(\boldsymbol{ABC}) = 2q - 1$$

341 在这种情况下,能够把(18)写为

$$0 \leqslant 3p - q \leqslant 2$$

而其他三个不等式只产生 $0 \leqslant p + q \leqslant 2$。设

$$x = P(\bar{a}bc) = P(a\bar{b}c) = P(ab\bar{c})$$

$$y = P(\bar{a}\bar{b}c) = P(\bar{a}b\bar{c}) = P(a\bar{b}\bar{c})$$

$$z = P(abc)$$

和

$$w = P(\bar{a}\bar{b}\bar{c})$$

不难表明,既然 $3x + 3y + z + w = 1$,所以,在边界 $3p = q$,根据不等式定义的值 $x = 0, y = q/3, z = 0, w = 1 - q$ 定义了一个可能的联合概率分布。在其他边界,$3p = q + 2$,这样一个可能的概率分布是,$x = (1 - q)/3, y = 0,\ z = q,\ w = 0$。于是,对于在这个不等式边界之内的 q 和 p 的任何值来说,我们能接受所选择的具有权重 $(3p - q)/2$ 和 $1 - (3p - q)/2$ 的这些分布的一个线性结合,使得权重概率增加到一,并获得联合概率分布:

$$x = \left(1 - \frac{3p - q}{2}\right)\frac{1 - q}{3}$$

$$y = \left(\frac{3p - q}{2}\right)\frac{q}{3}$$

$$z = \left(1 - \frac{3p - q}{2}\right)q$$

$$w = \left(\frac{3p - q}{2}\right)(1 - q)$$

这证明,如果满足这个不等式,就存在一个联合概率分布,因此,也存在一个非语境的隐变量,这样完成了证明。推广到不对称情况是冗长的但简单。

作为上面的不等式的一个结论,人们能够表明,在 GHZ 态中出现的关联可能是如此之强,以至于即使我们允许有实验误差,也仍然能证实,不存在一个联合分布(de Barros and Suppes, 2000)。

推论 1. 设 $\mathbf{A}$、$\mathbf{B}$ 和 $\mathbf{C}$ 是三个 ± 1 的随机变量,使得

(i) $E(\mathbf{A}) = E(\mathbf{B}) = E(\mathbf{C}) \geqslant 1 - \epsilon$

(ii) $E(\mathbf{ABC}) \leqslant -1 + \epsilon$

其中,ϵ 表示由于实验误差所观察的 GHZ 关联的减小。于是,不可能存在 $\mathbf{A}$、$\mathbf{B}$ 和 $\mathbf{C}$ 的一个联合概率分布,如果

$$\epsilon < \frac{1}{2}$$

证明: 为了看出这一点,让我们计算上面定义的 F 的值。我们立即得到

$$F = 3(1 - \epsilon) - (-1 + \epsilon)$$

但是,如果 $F \leqslant 2$,那么所观察的关联只是与一个非语境的隐变

342 量理论相一致，因此，$\epsilon < \frac{1}{2}$。于是，不可能存在满足(i)(ii)的 **A**、**B** 和 **C** 的一个联合概率分布，如果

$$\epsilon < \frac{1}{2} \tag{23}$$

从上面得到的不等式来看，显然，得到 GHZ 型的强于 0.5 的关联的任何实验，都不可能有一个联合概率分布。例如，最近在因斯布鲁克(Innsbruck, Bouwmeester et al., 1999)所做的三光子纠缠态的实验支持了量子力学的结果：根本不存在非语境的隐变量来说明这些态的关联。因此，根据对 GHZ 定理的重新表述，可能用强的仍然是不完善的实验关联来证明，非语境的隐变量理论与实验结果不一致。

二阶高斯定理。关于连续随机变量的有穷序列的一个基本的二阶表征定理如下：

定理 9. *设已知 n 个连续的随机变量，设它们的平均值、方差和协方差都存在而且是有穷的，同时，所有的方差都不为零。于是，与已知的平均值、方差和协方差相一致的 n 个随机变量的一个联合高斯概率分布存在的充分必要条件是，关联矩阵的本征值是非负的。*

这个定理的详尽讨论和证明能够在勒夫(Loève)的著作(1978)中找到。重要的是注意，这个定理的假设是，每对随机变量都足以假定，存在着一个惟一的二元变量的高斯分布，这个分布具有已知的一对平均值与一对方差，以及一对协方差。此外，正如所有 n 个变量的联合分布所要求的那样，如果关联矩阵的本征值都是非负的，那么存在着 n 个随机变量的惟一的高斯联合分布。

当有些关联或协方差是未知的时，为了把像贝尔那样的不

等式包括在内，我们阐述下一个表征定理。

定理 10. 设已知 n 个连续的随机变量，使得它们满足定理 4 的定域性假设，设它们的平均值和方差是存在的和有穷的，所有的方差都不等于零，再设 $m \leqslant n(n-1)/2$ 的协方差是已知的和有穷的。于是，下列两个条件是等价的。

(i) 存在着与已知的平均值、方差和协方差相一致的 n 个随机变量的一个联合概率分布。

(ii) 已知 $m \leqslant n(n-1)/2$ 的协方差，存在着可以赋值于缺失关联的实数，以便完备的关联矩阵具有的本征值全是非负的。

此外，(i)或(ii)意味着存在满足定域性条件Ⅱ和二阶因式分解条件的一个隐变量 $\boldsymbol{\lambda}$。

定理 10 的证明直接从定理 9 中得出。

运用定理 9，我们也能推出一个非线性的不等式，这个不等式是拥有一个联合分布的三个高斯随机变量的充分必要条件。 343
在这个表征定理的陈述中，$\rho(\boldsymbol{X}, \boldsymbol{Y})$是 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的关联。

定理 11. 设 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 是三个高斯随机变量，它们的平均值、方差和协方差都是已知的，而且，它们的方差不等于零。于是，存在着与已知的平均值、方差和协方差相一致的 $\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 的一个联合高斯分布(必定是惟一的)，当且仅当，

$$\rho(\boldsymbol{XY})^2 + \rho(\boldsymbol{XZ})^2 + \rho(\boldsymbol{YZ})^2 \leqslant 2\rho(\boldsymbol{XY})\rho(\boldsymbol{YZ})\rho(\boldsymbol{XZ}) + 1$$

证明可以根据关联矩阵的行列式直接得出。对于肯定是非负的一个矩阵来说，整个矩阵的行列式和所有的主子式一定大于等于零，

$$Det\begin{pmatrix} 1 & \rho(\boldsymbol{XY}) & \rho(\boldsymbol{XZ}) \\ \rho(\boldsymbol{XY}) & 1 & \rho(\boldsymbol{YZ}) \\ \rho(\boldsymbol{XZ}) & \rho(\boldsymbol{YZ}) & 1 \end{pmatrix} \geqslant 0 \tag{24}$$

包括我们有的子式条件在内,

$$\rho(\mathbf{XY})^2+\rho(\mathbf{XZ})^2+\rho(\mathbf{YZ})^2-2\rho(\mathbf{XY})\rho(\mathbf{XZ})\rho(\mathbf{YZ})\leqslant 1$$

$$\rho(\mathbf{XY})^2\leqslant 1$$

$$\rho(\mathbf{YZ})^2\leqslant 1 \tag{25}$$

后面两个不等式是自动地被满足的,因为这些关联以±1为界。

同时观察和联合分布。当观察是同时的和环境是稳定的与不变的时,结果是,根据可重复的同时观察,我们能够获得令人满意的频率数据,因此,存在着表示这些同时观察的所有随机变量的一种联合分布。注意,我们因而从上述所能得出的结论是:在所有这些情况下,由于存在着联合概率分布,所以,一定存在着一个因子隐变量。只根据这种考虑就能得出,违反贝尔不等式或其他隐变量标准的任何一个量子力学的事例,都必须是这样的:正在讨论的观察不能全部同时进行。把同时性标准推广到一个令人满意的相对论的标准是简单的。

7.3 因果性过程的弱可逆性和强可逆性①

众所周知并通常得到评论的是,经典物理学的基本方程,在从时间 t 变换到 $-t$ 的条件下,仍然是有效的。通常也评论说,在时间方向改变的条件下的这种不变性完全与日常经验相反。因此,很好地建立起来的经典物理学,还有物理学的其他部分,例如,相对论力学和量子力学中的薛定谔方程的不变性,产生了

① 本节的大多数资料也能从苏佩斯的著作(2001b)中找到。

关于因果性过程本性的一个自然的哲学忧虑。本节的目的不是以任何完备的方式解决物理学与日常经验之间的紧张关系，而至少是通过引入两个可逆性概念缓解这种紧张。第一个可逆性 344 概念是**弱可逆性**概念，在时间反向的条件下，经典物理学方程的不变性例示了这种情况。这是可逆性的**弱的意义**，因为我们通常能够通过观察辨别，一个质点力学系统是以哪种方式进行的。当我们转到一个反向的时间时，现象却是不相同。基本的例子是，当一个粒子从静止开始沿直线加速运动，并撞到一个不可穿透的墙上时，测量它的速度。特别是，速度改变的图像，在时间反转条件下，将完全不同。在这样的反转条件下，在运动开始时，粒子直接有一个很大的速度，然后，继续减速直到最后达到静止状态，与在通常的时间方向上所观察到的情况完全相反。当然，当在时间反转条件下讨论经典物理学的不变性时，人们完全承认这种不同。

这里，在对大量自然现象（包括随机过程）的不同分析中，提出和运用的**强可逆性**概念，具有更强的可逆性条件。如果我们不能区分一个过程（不管是决定论的过程，还是随机过程）是前进，还是后退，就把这个过程说成是强可逆的。用通常使用的更生动的术语来说，观看影片倒转的过程，也就是，反过来放电影的过程，看不出与影片正放有什么区别。在下文中，这些思想会变得更精确，特别是对于随机过程。我应该提到，这里所用的强可逆性概念对应于随机过程中通常的简单**可逆**。因此，下面给出的是强可逆的马尔可夫链的条件，与已有的标准概念，例如，费勒（Feller）著作中的马尔可夫链的标准概念，是相同的。

尽管下文中提出的许多形式上的区分，在概率理论特别是随机过程的文献中，是熟悉的，但在我的印象中，在时间可逆性的哲学讨论中，这里提出的区分，作为一种区分，没有得到充分

的强调。在不是强可逆的日常经验和类似的物理学中理想质点的许多轨道或其他重要过程之间,存在着一种密切的相似,在这种意义上,引入这种区分,是为了缓解我一开始提到的那种紧张关系。

弱可逆性。正如所评论的那样,众所周知,经典质点力学和相对论的质点力学具有弱可逆的特性,即改变时间方向的变换总是从经典质点力学系统带到经典质点力学系统,而且,相对论的情形也是如此。这些结果的详细证明可以在麦金西与苏佩斯(1953)关于经典质点力学的著作中和鲁宾(Rubin)与苏佩斯(1954)关于相对论的质点力学的著作中找到。

证明把一阶马尔可夫链带到时间反向条件下的一阶马尔可夫链是简单的。下面是对一个有穷状态数的一阶马尔可夫链的基本证明。在这个证明中假设,表达式的分母中出现的概率都大于零。

345 **定理 1.** *所有的一阶马尔可夫链都是弱可逆的。*

证明:(针对有穷态的离散时间过程)

$$P(i_{n-1} \mid j_n k_{n+1}) = \frac{P(i_{n-1} j_n k_{n+1})}{P(k_{n+1} \mid j_n) P(j_n)}$$

$$= \frac{P(k_{n+1} \mid j_n) P(j_n \mid i_{n-1}) P(i_{n-1})}{P(k_{n+1} \mid j_n) P(j_n)}$$

$$= \frac{P(j_n i_n - 1)}{P(j_n)}$$

$$= P(i_{n-1} \mid j_n)$$

量子力学中的弱可逆性情形,要稍微更复杂些。当然,薛定谔方程也是如此可逆的。在改变时间方向的变换条件下,薛定谔方程是弱可逆的。另一方面,在标准的解释中,测量过程并非

如此(von Neumann 1932/1955, Ch. 5)。

最后,从常识的观点来看,日常经验肯定不是弱可逆的。时间有固定的方向,而且,根据这个方向的人类体验的经验证据是无法抵抗的。

$$\text{向前} \neq \text{向后}$$

人们几乎从来不会倒着上楼梯。没有一种赛跑是向后跑,等等。这不是主张"影片倒转"违反了力学规律,违反了仅有的归纳律和经验事实,最特别的是大多数被普遍接受的人类记忆的本性:

$$\text{未知的未来} \neq \text{已知的过去}$$

某些定义。在进一步阐述之前,引入定理 1 中潜在地假设的某些标准的随机概念,是令人向往的。首先,设 $\mathbf{X}(t)$是一个随机过程,使得对于每个时间 t, $\mathbf{X}(t)$在一个有穷集中取值,这只是为阐明的简单性强加的一种限制。任意有穷族 $\mathbf{X}(t_1)$, $\mathbf{X}(t_2)$, …, $\mathbf{X}(t_n)$有一个联合概率分布。用明显的简化符号表示,一个过程是(一阶)**马尔可夫过程**,当且仅当,对于任意时间 $t_1 < t_2$, …, t_n,

$$P(x_n \mid x_n - 1, \cdots, x_1) = P(x_n \mid x_{n-1}) \tag{1}$$

其中,把简化的符号定义为

$$P(x_n \mid x_{n-1}) = P(\mathbf{X}(t_n) = j_n \mid \mathbf{X}(t_{n-1}) = i_{n-1})$$

如果像(1)那样的跃迁概率与时间无关,那么一个马尔可夫过程是**齐次的**(homogeneous),如果在有限的几步之内,能够从任意其他态到达任意态 j,那么它是**不可还原的**。对于时间离散的齐次马尔可夫过程,我用几个不同的但有用的符号写出跃迁概率

$$P_{ij} = p_{i,j} = P(j_n \mid i_{n-1}) = P(\mathbf{X}_n = j \mid \mathbf{X}_{n-1} = i)$$
$$= P(\mathbf{X}(t_n) = j_n \mid \mathbf{X}(t_{n-1}) = i_{n-1})$$

在概率文献中，所有这些都是常用的。

*设 $P_{ii}(n)$是在 n 次跃迁中返回到态 i 的概率。一个态 i 的*周期$r(i)$是整数 n 的集合的最大公约数，其中，$P_{ii}(n) > 0$。一个态是**非周期性的**，如果它有周期 1。此外，不难表明，如果一个马尔可夫过程是不可还原的，那么每个态都有相同的周期。因此，一个非周期的过程是每个态都有周期 1 的过程。

346

此后，当我说到一个马尔可夫**链**时，我假设了一个有穷态的、离散时间的、同类的、不能还原的和非周期的马尔可夫过程。

如果不加改变运用到连续时间过程，是马尔可夫的、同类的和不可还原的定义全都是可应用的，但是，为了避免物理上不可能的情况，强加了一些进一步的条件。首先，我们要求这样一个过程在任何一个态都保持一个正的时长，第二，这个过程在有限的时间内不可能经过一个无穷序列的态。与离散时间过程的跃迁概率相对应，连续时间过程的**跃迁率**定义为

$$q_{ij} = \lim_{\tau\to 0} \frac{P(X(t+\tau) = j \mid X(t) = i)}{\tau}, \ i \neq j,$$

而且，为了达到定义的目标，我们使 $q_{ii} = 0$。此后，当我说到一个马尔可夫**过程**时，我将意指一个连续时间过程。在每个态都保持这样一个过程，因为一段时间的指数分布具有的参数 $q(i)$ 定义为

$$q(i) = \sum_j q_{ij}$$

而且，离开态 i 时，跃迁到态 j 的概率是

$$p_{ij} = \frac{q_{ij}}{q_i}$$

注意在一般情况下，

$$\sum_j q_{ij} \neq 1$$

即从态 i 到另一个态的跃迁率不需要增加到 1,时间离散过程中从态 i 的跃迁概率一定也是这种情况。[这样的连续时间过程的一个好的讨论可以在凯利(Kelly)的著作(1979)中找到。]

强可逆性。强可逆性的直观思想很容易根据电影来描述。如果一位观众无法说出一个影片是正着放还是倒着放,这个影片就是强可逆的。在这两个方向上,根本没有感知上的差别。现在,正如我们看电影的日常经验直接告诉我们的那样,这完全是一个例外的情形。另一方面,存在着确实具有强可逆性的重要的物理过程,我们将讨论其中的一些过程。首先一个有趣的情况是这样的。一个马尔可夫链什么时候是强可逆的呢?对此的正确背景是,我们立刻限于平稳的、各态历经的马尔可夫链。一个马尔可夫链是**平稳的**,如果它的平均分布总是相同的。如果它有一个独立于这个初始分布的一个惟一的渐近分布,那么它是**各态历经的**。注意,在这些定义条件下,一个马尔可夫链,如果它的初始分布不同于它的惟一的渐近分布,它可能是各态历经的,但不是平稳的。但我在这里假设,各态历经链是平稳的。

一个各态历经的马尔可夫链是**强可逆的**,如果

$$\pi_i p_{ij} = \pi_j p_{ji} \qquad (2)$$

对于所有的态 i 和 j,这里,π 是惟一的渐近密度。方程(2)通常被称为强可逆性的**细致平衡**条件。伯努利过程,比如,掷硬币,是强可逆的过程,因为没有过去态的记忆:$p_{ij} = p_j$ 和 $p_{ji} = p_i$,因此,我们满足(2)。此外,任意 2-态的各态历经的马尔可夫链是强可逆的。这里是证明。既然这个过程是各态历经的, 347

$$\pi_1 = \pi_1 p_{11} + \pi_2 p_{21} = \pi_1(1 - p_{12}) + \pi_2 p_{21}$$

因此，

$$\pi_1 p_{12} = \pi_2 p_{21}$$

另一方面，许多 3 -态链不是强可逆的。这里是一个三色旋转器的例子。

	B	R	Y
B	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
R	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
Y	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$

我们立即有三态的平均密度 π：

$$\pi_b = \frac{1}{2}\pi_b + \frac{3}{8}\pi_r + \frac{1}{8}(1 - \pi_b - \pi_r)$$

$$\pi_r = \frac{1}{4}\pi_b + \frac{1}{2}\pi_r + \frac{3}{8}(1 - \pi_b - \pi_r)$$

解这两个方程，我们发现，

$$\pi_b = \frac{13}{37},\ \pi_r = \frac{14}{37},\ \pi_y = \frac{10}{37}$$

这里是它不是强可逆的证明：

$$\pi_b p_{br} = \frac{13}{17} \cdot \frac{1}{4} \neq \frac{14}{37} \cdot \frac{3}{8} = \pi_r p_{rb}$$

我们能把我们关于马尔可夫链所说的概括为下列定理，其中，在这个定理的符号中，$\overset{d}{=}$ 意指相同的分布。

定理 2. 如果一个各态历经的马尔可夫链是强可逆的，那么

不可能区分一个影片是正放还是倒放，也就是说，当随机变量的顺序颠倒时，这个过程的随机变量的任何一个有穷序列都有相同的分布：

$$(\boldsymbol{X}_0, \boldsymbol{X}_1, \cdots, \boldsymbol{X}_n) \overset{d}{=} (\boldsymbol{X}_n, \boldsymbol{X}_{n-1}, \cdots, \boldsymbol{X}_0)$$

我给出 $n = 2$ 的证明：

$$\begin{aligned} P(\boldsymbol{X}_0 = i, \boldsymbol{X}_1 = j) &= P(\boldsymbol{X}_1 = j \mid \boldsymbol{X}_0 = i) P(\boldsymbol{X}_0 = i) \\ &= \pi_i p_{ij} \\ &= \pi_j p_{ji} \qquad \text{根据(2)} \\ &= P(\boldsymbol{X}_1 = i \mid \boldsymbol{X}_0 = j) P(\boldsymbol{X}_0 = j) \\ &= P(\boldsymbol{X}_1 = i, \boldsymbol{X}_0 = j). \end{aligned}$$

正如上面定义的那样，通过用跃迁率 q_{ij}（其中，$q_{ij} \geqslant 0, i \neq j$）替代跃迁概率，改变了所引入的具有有穷态连续时间的概念。 348
定义强可逆性的细致平衡条件在形式上与 $\pi_i \geqslant 0$ 的(2)的那些条件相同，而且，与前面一样，求和等于 1：

$$\pi_i q_{ij} = \pi_j q_{ji} \tag{3}$$

根据恰好像证明定理 2 一样的证明，我们直接有

定理 3. *如果一个各态历经的、时间连续的马尔可夫过程是强可逆的，那么不可能区分一个影片是正放还是倒放。*

在当前的语境中，关于一阶马尔可夫链没有什么特殊的。一个二阶各态历经的马尔可夫链是强可逆的，当且仅当，

$$\pi_i \pi_{ij} p_{ijk} = \pi_k \pi_{kj} p_{kji} \tag{4}$$

在一个链是二阶马尔可夫链的地方，如果

$$P(x_n \mid x_{n-1}, \cdots, x_1) = P(x_n \mid x_{n-1}, x_{n-2})$$

对应于一阶的(1)。

定理 4. 如果一个二阶马尔可夫链是各态历经的和强可逆的，那么不可能区分一个影片是正放还是倒放，即

$$(\boldsymbol{X}_0, \boldsymbol{X}_1, \cdots, \boldsymbol{X}_n) \stackrel{d}{=} (\boldsymbol{X}_n, \boldsymbol{X}_{n-1}, \cdots \boldsymbol{X}_0)$$

证明： 当 $n = 3$ 时，

$$P(\boldsymbol{X}_0 = i, \boldsymbol{X}_1 = j, \boldsymbol{X}_2 = k)$$

$$= P(\boldsymbol{X}_2 = k \mid \boldsymbol{X}_1 = j, \boldsymbol{X}_0 = i)P(\boldsymbol{X}_1 = j \mid \boldsymbol{X}_0 = i)P(\boldsymbol{X}_0 = i)$$

$$= p_{ijk}\pi_{ij}\pi_i$$

$$= \pi_k\pi_{kj}p_{kji} \quad \text{根据(4)}$$

$$= P(\boldsymbol{X}_2 = i \mid \boldsymbol{X}_1 = j, \boldsymbol{X}_0 = k)P(\boldsymbol{X}_1 = j \mid \boldsymbol{X}_0 = k)P(\boldsymbol{X}_0 = k)$$

$$= P(\boldsymbol{X}_2 = i, \boldsymbol{X}_1 = j, \boldsymbol{X}_0 = k)$$

显然，这个证明也适用于连续时间的二阶马尔可夫过程。毫不费力就能把这个同类证明推广到各态历经的和平稳的无穷阶的链。在兰佩蒂(Lamperti)和苏佩斯的文章中(1959)提出了所要求的推理思路。

当再次从马尔可夫链转向连续时间的马尔可夫过程时，强可逆过程的一个简单例子是各态历经的**生死过程**，这些过程是通过跃迁率等于零来定义的，除了 $q(i, i+1) > 0$ 代表生和 $q(i, i-1) > 0$ 代表死之外。因此，细致平衡条件(3)假设了这种形式：

$$\pi_i q_{i,\, i+1} = \pi_{i+1} q_{i+1,\, i} \tag{5}$$

其中，π_i 成为态的有穷集合的平稳概率密度。

埃伦费斯特(Ehrenfest)模型。 统计力学的一个简化模型漂亮地例示了一个强可逆的生死过程。在麦克斯韦的影响下，存在着两个粒子容器，认为充满了理想气体分子。设 $\mathbf{X}(t)$ 是容器

349

Ⅰ中的粒子数,因此,$k-\mathbf{X}(t)$ 是容器Ⅱ中的粒子数。“生与死”的过程对应于容器Ⅰ中的粒子数的增减,具有的跃迁率是

$$q_{i,\ i-1} = i\lambda, \qquad i = 1, 2, \cdots, k,$$

$$q_{i,\ i+1} = (k-i)\lambda, \quad i = 0, 1, \cdots, k-1,$$

其中,λ 是比率参数。可以表明,平稳密度是

$$\pi_i = 2^{-k}\binom{k}{i} \qquad i = 0, 1, \cdots, k,$$

其中,

$$\binom{k}{i} = \frac{k!}{(k-i)!i!} \text{and } 0! = 1$$

通过检查细致平衡条件(5)可以看出,这个过程是强可逆的。

$$\pi_i q_{i,\ i+1} = 2^{-k}\binom{k}{i}(k-i)\lambda$$

$$= \frac{2^{-k}k!(k-i)\lambda}{(k-i)!i!}$$

$$= \frac{2^{-k}k!\lambda}{(k-(i+1))!i!}$$

$$\text{既然}\frac{k-i}{(k-i)!} = \frac{1}{(k-(i+1))!}$$

$$= \frac{2^{-k}k!(i+1)\lambda}{(k-(i+1))!(i+1)!} \quad \text{既然}\frac{1}{i!} = \frac{i+1}{(i+1)!}$$

$$= 2^{-k}\binom{k}{i+1}(i+1)\lambda$$

$$= \pi_{i+1} q_{i+1,\ i}$$

不用企图详细地讨论这个过程的熵,我注意到,即使它是平

稳的和强可逆的，但在这两个容器中，从一个瞬间到另一个瞬间的平均涨落，在两个容器中，总是更强地趋向于粒子的相等分布，而不是远离这种相等分布，对于

$$q_{i,\ i-1} = i\lambda < (k-i)\lambda = q_{i,\ i+1}$$

当且仅当，$i < \frac{k}{2}$，独立于比率参数 λ。如果我们根据容器 I 中的粒子数在时间 t 的相对频率 $f(t)$来计算熵，那么这个过程在 t 的瞬时熵是

$$H(t) = -(f(t)\log f(t)) + (1-f(t))\log(1-f(t))$$

而且，根据刚才给出的分析，熵的平均变化率总是在增加，即是正的，尽管这个过程是强可逆的。但显然，即使在平衡态，即当平稳时，这个模型的这两个特征之间并没有矛盾。对埃伦费斯特模型，还有玻尔兹曼方程，以及洛施密特和策梅洛的联合"悖论"的一个出色的详细分析，可以在卡克(Kac)的著作(1959)中找到。

决定论的系统。我现在转向决定论的经典物理学系统。应该很明显的是，只有很特殊的系统是强可逆的。任何类型的耗散系统一般情况下都不是强可逆的。打台球就不是强可逆的。350 我们打出球，由于桌子的摩擦，球会在某个地方停下来。我们能够轻易区分出正着放和倒着放的影片。另一方面，在一个理想的台球桌上打理想的台球，如果总能量守恒，入射角和反射角令人满意地严格相等，打出去的球可能是做周期性的运动，而且，一旦运动起来，将永远继续下去。此外，这种理想运动当然将是强可逆的。通过"观看"这种运动，我们不能决定我们看到是正着放影片还是倒着放影片。(我给"观看"加上引号，因为事实上没有这样的台球情形，但我们能在一个有限的时间周期内进行

很好的人工模拟。)

具有强可逆性的经典系统的一个简单的但重要的例子,是一维谐振子的无阻尼和无驱动的振动。这样一个谐振子的微分方程是

$$\frac{d^2x}{dt^2}+\omega^2x=0$$

其中,ω 是当无阻尼和无驱动时这个谐振子的自然频率,而且,在时间 $t=0$ 时的初始条件是

$$x_0=\alpha,$$

$$\frac{dx_0}{dt}=0。$$

(当这个模型是单摆时,相应的初始条件是,在时间 $t=0$ 时,单摆静止,具有的位移是 α。)(1)的一般解是

$$x(t)=A\cos\omega t+B\sin\omega t$$

和

$$\frac{dx}{dt}=-A\omega\sin\omega t+B\omega\cos\omega t$$

因此,在 $t=0$ 时, $B=0$, 因而,

$$在\ t=0\ 和\ A=\alpha\ 时,A\cos\omega t=\alpha$$

于是,

$$\frac{d^2x}{dt}=-A\omega^2\cos\omega t$$

而且,最终,

$$x(t)=A\cos\omega t \tag{6}$$

我们立刻有,这样一个谐振子是强可逆的,因为对于所有的 t

$$x(-t) = A\cos -\omega t = A\cos\omega t = x(t)$$

注意,这个解释对于 $-\infty < t < \infty$ 成立,而且,在 $t = 0$ 时的条件实际上不是初始条件,而只是在形式上产生一个简单解的某个 t 的条件。

另一方面,阻尼的、无驱动的谐振子的微分方程是

$$\frac{d^2x}{dt^2} + 2k\frac{dx}{dt} + \omega^2 x = 0$$

其中,k 是阻尼系数,很容易表明,这个谐振子不是强可逆的。这种结果不足为奇,因为一个阻尼谐振子(一维的)是一个耗散系统的最简单的物理学例子。在标准情况下,在阻尼不太强的

351 地方,振动函数的振幅随时间增加。更明确地说,振子的包络是负指数形式:正面为 e^{-kt},负面为 $-e^{-kt}$。

无疑,如果目标是找到在经典物理学中不是强可逆的系统的例子,因而因果性分析不是强可逆的,那么就在许多类型的耗散系统中去寻找。在自然界中最普遍的这些系统强化了我们的个人经验,支持常识的观点,正如谚语所说,"……当然,时间是不可逆的。谁会想到其他方面呢?"

8. 语言的表征

本章给出关于语言(例如,语境无关的语言)的各种表征。353
第一种表征是关于不同长度的自动机方面的,而且,这种表征以两种方式进行,即根据自动机的语言和通过语言表征的自动机。一旦提出如何剖析或编码一种语言的问题,这两种表征就如此密切地交织在一起,这一事实也许是令人吃惊的,但却是可理解的。本章的前两节讨论这些问题。现在的每一位年轻的计算机科学家都会接受这方面内容的教育。在 20 世纪 70 年代的许多年里,我把这些内容作为自动机理论课的一部分来讲授。

下面两部分与我自己的早期工作有着更特殊和直接的关系。第 3 节主要表明,强化和条件作用的简单抽象的思想,在严格公理化时,如何充分地说明,刺激-反应的学习结构能表示任何一个有穷自动机,因而自动机的正则语言(regular language)能表示任何一个有穷自动机。这一节的最后一部分经过无限寄存器机的刺激-反应表征,把这个表征推广到任何一个可计算的函数。也许事实是,大多数语言学家完全不相信,联想或条件作用的简单机制足以说明复杂而微妙的语言现象。这一节要证明的内容在数学上太抽象了,无法说服这些语言学家中的许多人,但这并不是所要完成的任务。本节的目标只是以一种严格的方

式表明,“原则上”如何能作出这样一种简化。本章的第 5 节和第 6 节包括了若干必要的更详细的工作。

第 4 节聚焦于上一节阐述的刺激-反应模型与只根据行为变量表述的学习模型的关系。主要定理是关于根据前面模型对后面模型的表征定理。冗长的证明表明了,即使在相对简单的情况下,证明把一种理论还原为另一种理论的详细的表征定理,也是困难的。一旦引入附加的理论概念,表征定理也对表明这样的问题有某种兴趣:即如何能用一阶马尔可夫过程表示依赖于遥远过去的当前事件的各种模型。例如,几乎人人都相信,孩提时代的事件对一位成年人的任何持续的影响,在这个人当前的身体与心理结构中,一定有体现,而且,当年的事件在事隔多年之后不是直接地、无中介地对人产生影响。但是,对于绝大多数情况来说,用现在可利用的数据和科学概念,详细地描述当前结构的具体化,是不可能的。这一节举出的例子表明,甚至在这种结构分析的最简单化的情形中,也必须作出这些考虑。

354

第 5 节致力于机器人的自然语言的机器学习。“机器学习”这个短语,事实上是指开发计算机学习的学习程序。这里把这样的学习结果表达为是理解那些具有联想语义学的文法。表征的相关意义是,学习如何用十种语言中的任何一种语言,比如,英语、汉语(普通话)等,表示简单的机器人命令的表征。这个案例研究可以说是为了表明,运用(广义地说)也适用于人类语言学习的原理,如何能学习一种自然语言的一个系统片段。然而,这里所举的机器学习的例子决不意味着模仿人类的语言学习,从这一节给出的细节来看明显如此。

第 6 节从学习转向概述我自己最近在语言和大脑方面的一些工作,这些工作是与较年轻的同事一起做的。焦点是用脑电图表示词和句子。当前工作的目标是试图正确地描绘,当人们

听到或看到某个词或句子之后,他的大脑是如何最直接地表达语言的。我们已经发现了某些有趣的不变性结果。这一节更一般的目标是在某种程度上给出,在这些研究中特别是在概率问题上,所用的方法论的详细意义。

这一章没有陈述或证明不变性定理。有些定理当然是可能的,但它们不在所考虑的概念和理论的主要发展范围之内。但第 6 节提供了某些有意义的经验的不变性,这些不变性是在对语言的神经处理过程中从实验上发现的。

第 7 节,即最后一节,是有关科学中的表征和还原的一个结语。这些评论主要是历史性的,并打算为本章前面和前几章的详细工作提供一个视角。

8.1 形式语言的层次结构

我们以当代语言学和计算机科学中广泛使用的形式语言的定义为出发点。这个层次结构的基本概念最初归功于乔姆斯基(1959)。我们开始于任意的短语结构文法的一般概念。通过增加限制,我们得到了其表达能力更有限的文法。

设 V 是一个有穷非空集。于是,V^* 是 V 的元素的所有有穷序列的集合。设$\varnothing$是空序列。于是,$V^+ = V^* - \{\varnothing\}$。$V$ 是词汇表。V^* 是字符串的集合。

定义 1. 一个结构 $G = (V, N, P, S)$ 是一种短语结构文法,当且仅当,

1. V、N 和 P 是非空有穷集,
2. $N \subseteq V$,
3. $P \subseteq V^+ \times V^*$,
4. $S \in N$。

355 标准术语是这样的。集合 N 是**非终结符**或**变量**的集合。$V_T = V - N$ 是**终结符**或词的集合。P 是产生式集合(the set of productions)。如果 $(\alpha, \beta) \in P$,我们通常写作:$\alpha \rightarrow \beta$,来表明,我们可以从 α"产生"β。最后,S 是**起始符**,派生由此开始。这些公理的弱化与短语结构文法定义的普遍性相一致。最感兴趣的大概是关注下面定义的各种受限制的类。

如果 γ 和 δ 在 V^* 中,那么 $\gamma\alpha\delta \Rightarrow G\, \gamma\beta\delta$,当且仅当,$\alpha \rightarrow \beta$ 是一种产生式,即 $(\alpha, \beta) \in P$。我们说,在这种情况下,字符串 $\gamma\beta\delta$ 是在文法 G 中从字符串 $\gamma\alpha\delta$ **直接派生出来的**。我们接下来在一种文法中根据直接派生来定义派生概念。从直观上看,派生是我们表明一个特定的字符串属于由该文法生成的语言的方式。

如果 γ_1、$\gamma_m \in V^*$,那么 $\gamma_1 \Rightarrow_G^* \gamma_m$,当且仅当,$\exists\, \gamma_2, \cdots, \gamma_{m-1} \in V^*$,使得 $\gamma_1 \Rightarrow G\, \gamma_2, \cdots, \gamma_{m-1} \Rightarrow G\, \gamma_m$。**通过一种文法 G 所产生的语言**[用 $L(G)$表示]是

$$L(G) = \{w \mid w \in V_T^* \ \&\ S \Rightarrow_G^* w\}$$

注意,对于这种语言中的一个字符串来说,它必须只是由终结词组成的,而且,它在 G 中的派生必须从起始符 S 开始。

我们说,两种文法 G_1 和 G_2 是**弱等价的**,当且仅当,它们生成相同的语言,即 $L(G_1) = L(G_2)$。注意,这没有更多地讨论表达结构(the structure of the utterances)。只是说在不关注结构的情况下,这些表达正好产生了相同的字符串集合。

形式语言的一个重要应用是分析自然语言的结构。确实,我们能声称,自然语言符合我们将要定义的层次结构。"这是否为真"这个相当复杂的问题,在这里实际上不会以任何令人满意的方式加以考察。当然证据很重要,用本节引入的文法概念能

够分析自然语言的重要片段。我们现在考虑一个简单的例子。

$N = \{S, NP, VP, PN, AdjP, Adj, CN, Art, IV, TV\}$

其中，S 是起始符，NP 表示名词短语，VP 表示动词短语，PN 表示专有名词，$AdjP$ 表示形容词短语，Adj 表示形容词，CN 表示普通名词，Art 表示冠词，IV 表示不及物动词，以及 TV 表示及物动词。

V_T = {你根据英语词汇的通常文法范畴来挑选终结英语词}

P = 下列产生式规则的集合，这里省略了产生终结词的词汇规则——一个词汇规则是它的右边项是一个终结词的规则：

1. $S \to NP + VP$
2. $NP \to PN$
3. $NP \to AdjP + CN$
4. $AdjP \to AdjP + Adj$
5. $AdjP \to Adj$
6. $AdjP \to Art$
7. $VP \to IV$
8. $VP \to TV + NP$

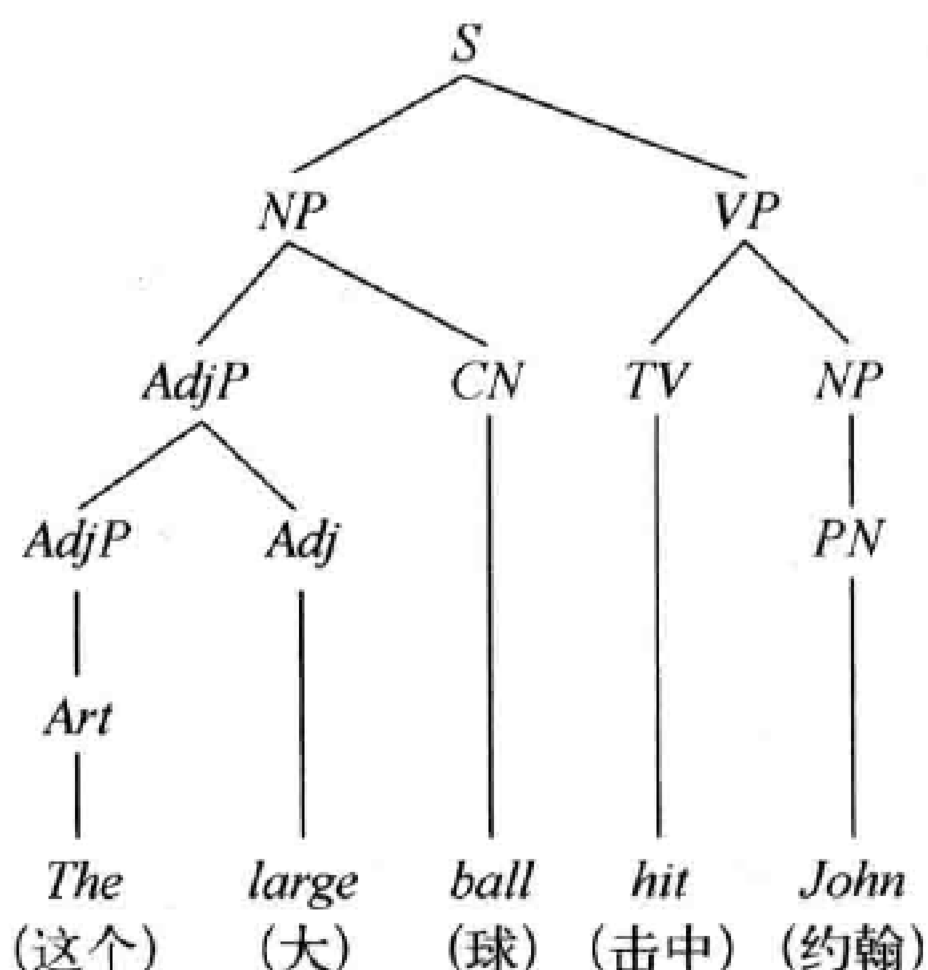
这些规则似乎应该是熟悉的，但当然是不完备的。在陈述这些规则时，为了避免在字符串的串接时造成混淆，所用的加法符号只表明串接。

下面是一种派生的例子。 356

1. S 起始符
2. $NP + VP$ 根据规则 1
3. $AdjP + CN + VP$ 根据规则 3
4. $AdjP + Adj + CN + VP$ 根据规则 4
5. $Art + Adj + CN + VP$ 根据规则 6
6. $Art + Adj + CN + TV + NP$ 根据规则 8
7. $Art + Adj + CN + TV + PN$ 根据规则 2
8. *The large ball hit John.*（这个大球击中约翰）

   根据词汇规则

在许多方面,派生树比线性派生更自然。刚才举的例子生成下列树。



我们将不从形式上定义派生树,尽管它们对于语境无关的文法(context-free grammars)是重要的。

文法类型。我们现在转向文法类型。0-型文法由定义1来描述。为了获得1-型或语境敏感的文法(context-sensitive grammars),我们增加限制:对于每个产生式 $\alpha\rightarrow\beta$, $|\alpha|\leqslant|\beta|$,其中,$|\alpha|$ 是 α 的长度,或者,α 中的符号数。这种限制能被等价地表示为

$$\alpha_1 A\alpha_2 \rightarrow \alpha_1\beta\alpha_2$$

其中,$\beta\neq\varnothing$ 且 $A\in N$。

为了获得2-型或语境无关的文法,我们增加更强的限制:如果 $\alpha\rightarrow\beta\in P$,那么

(i) α 是一个变量,即 $\alpha\in N$,因此,产生式具有的形式是 $A\rightarrow\beta$,

(ii) $\beta\neq\varnothing$。

为了获得3-型或正则文法(regular grammars),我们增加

还要更强的限制：任何一种产生式必须具有的形式是

(i) $A \to aB$，或者，

(ii) $A \to a$，

其中，$A, B \in N, a \in V_T$。[我们能用(i′)$A \to Ba$ 替代(i)，但从右线性变化到左线性是不重要的，即使必须改变形式(i)的所有规则。]

注意，我们能把(i)写成 $A \to a + B$ 表明了在写个别规则时表示串接的加法符号的熟悉用法。语境敏感文法、语境无关文法和正则文法的定义是，它们生成的语言排除了空字符串，0-型文法并非如此。为了达到某些技术上的目的，所希望的不是作出这种排除，但这里只是为了简化解释才这样做。 357

就一个正则文法的例子而言，我们可以通过使 V_T 只是某些词的文法范畴的符号，稍微改变一下我们前面的例子。

$$N = \{S, NP, VP, AdjP\}$$

$$V_T = \{PN, CN, Adj, Art, IV, TV\}$$

如果我们不再讨论规则 8，并把规则 1 重写为 1′，$S \to NP + IV$，我们能直接核实，这种文法是正则的。重写规则 8 更棘手，而且要求更多的变量，即 N 个元素。我只概述这种抨击。我们增加像下面那样的规则：

$$S \to X + CN$$

$$X \to Y + Adj$$

$$X \to Y + Art$$

$$Y \to VP' + Adj$$

$$Y \to VP' + Art$$

$$VP' \to TV$$

一般情况下,很难看出,一种语言是否有一种已知类型的文法。

当在公理意义上用集合论谓词定义理论时,不同类型的短语结构文法定义的抽象性,典型地代表了各个学科中的定义情况。不对在词表等集合中可能出现的对象类型作出进一步的限制,就能够构造出直观上很奇怪的文法——奇怪,不是因为它们的结构特征,而是因为它们的内容特征。

本节考虑了关于所定义的文法类型的一些重要定理。我首先陈述关于一种特定语言的文法范式(normal forms)的某些结果,然后,陈述关于语言的集合论运算的某些定理,接下来陈述关于不可解问题和模糊问题的一个简单结果。本节最终触及的主题是,语境敏感语言的一种重要的潜在自然语言中的应用,这种语境敏感的语言是用词汇功能文法来定义的。

我在本章不打算给出像形式语言的详尽发展那样的阐述,只打算给出针对这些理论结构所研究的问题类型的意义。

范式。我从关于语境敏感语言和 0 -型语言的定理开始。注意,比如说,如果至少存在着生成一种语言的语境敏感的文法,那么这种语言是语境敏感的。

定理 1. 每一种语境敏感的语言都能够通过这样一种文法生成:在这种文法中,所有的产生式都具有形式 $\alpha \to \beta$,其中,α 和 β 是含变量的字符串,[1]或者,具有形式 $A \to b$,其中,A 是一个变量,b 是一个终结词。此外,每一种 0 -型语言都能够通过这样一种文法生成:它的产生式具有两种相同的形式。

358

证明: 设 $G = (V, N, P, S)$ 是一种语境敏感的文法。对于 V_T 中的每个 a,设 X_a 是一个不在 N 中的新符号。现在定义

---

① 这里的符号与原文不同,这是作者的学生王瑞博士在帮着校译本节时,经过与作者沟通后修改的,下文中出现的相同符号也做了相应的修改。——译者

文法 $G' = (V', N', P', S)$，其中，

$$N' = N \cup \{X_a \mid a \in V_T\}$$

$$V' = V \cup N'$$

而且，P'的定义如下。首先，对于 $a \in V_T$ 来说，每个形式为 $X_a \to a$ 的产生式都在 P'中。其次，如果 $\alpha \to \beta$ 在 P 中，那么$\alpha_1 \to \beta_1$ 在 P 中，其中，α_1 是通过用 X_a 代替 V_T 中每次出现的 a，从 α 得到的，β_1 是以同样的方式从 β 得到的。通过这种构造，不难看出，$L(G') = L(G)$。同样明显的是，如果 G 是 0-型文法，那么证明是类似的。

这种证明的简单性是关于形式语言的许多有趣的定理的典型特征。整个问题是定义正确的新的实体——在当前的情况下是定义文法 G'。于是，证明所定义的实体具有适当的特性，通常是简单明了的，但如果很详细地完成的话，经常是单调乏味的。在几乎所有的情况下，这种证明含蓄地描述了构造这种新的文法或在该定理中涉及的其他实体的一种算法。

定理 2. (乔姆斯基的范式，1959) *任何一种语境无关的语言都能通过这样一种文法生成：在这种文法中，所有的产生式都具有形式 $A \to BC$ 或 $A \to a$，其中，A、B 和 C 是变量，a 是一个终结词。*

证明： 设 $G = (V, N, P, S)$ 是生成语境无关语言的一种语境无关的文法。从 G 我们构造一种新的文法 G'满足这个定理的必要条件。我们首先描述 P'：

(i) 把 P 中形式为 $A \to BC$(其中，A、B、C 在 N 中)的每个产生式都放入 P'中，

(ii) 把 P 中形式为 $A \to a$(其中，A 是在 N 中，a 在 V_T 中)的每个产生式都放入 P'中，

(iii) 对于 P 中每个产生式 $A \to X_1 \cdots X_k$（其中，$k > 2$，A，X_1，…，X_k 在 V 中），把下列产生式放入 P' 中：

$$A \to X'_1 Y_2$$

$$Y_2 \to X'_2 Y_3$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$Y_{k-2} \to X'_{k-2} Y_{k-1}$$

$$Y_{k-1} \to X'_{k-1} X'_k$$

其中，每个 Y_j 都是不在 N 中的一个新变量；如果 X_i 在 N 中，那么每个 X'_i 都是 X_i；如果 X_i 在 V_T 中，那么每个 X'_i 都是不在 N 中的一个新变量。

(iv) 对于形式为 $A \to X_1 X_2$（其中，X_1 或 X_2 或两者都在 V_T 中）的每个产生式来说，把产生式 $A \to X'_1 X'_2$ 放入 P' 中。

(v) 对于(iii)或(iv)中属于 V_T 的任意 X_i 来说，把产生式 $X'_i \to X_i$ 放入 P' 中。

359 为了完成 G' 的构造，设 N' 是 N 和在构造的 P 中引入的所有的新变量，设 $V' = V \cup N'$ 和 $S' = S$。$L(G') = L(G)$ 的明显证明留给读者。

证毕。

我们不加证明地陈述语境无关语言的另一个著名的范式。

定理 3. [格雷巴赫(Greibach)的范式，1965] **每一种语境无关的语言都能够通过这样一种文法生成：在这种文法中，每个产生式都具有形式 $A \to a\beta$，其中，A 是一个变量，a 是一个终结词，β 是一个可能为空的变量字符串。**

像在有关形式系统的其他研究情形中那样，我们所考虑的范式，

在简化和标准化语境无关的语言或其他语言所用的符号时，是有用的。像在命题逻辑中的析取范式和合取范式的情形中那样，这些形式在证明语言的各种特性时也起到了直接的作用。例如，乔姆斯基的语境无关文法的范式被直接用来证明，存在着一种算法来确定，一个已知的语境无关的文法 G 是生成了一种有穷语言，还是一种无穷语言，即确定 $L(G)$ 是一个有穷集，还是一个无穷集。

关于语言的运算。设 G 和 G' 是两种短语结构文法。存在着关于语言 $L(G)$ 和 $L(G')$（特别是作为集合考虑时）的各种自然运算。于是产生了哪一类语言在这些运算条件下是闭合的问题。并集和交集显然都是运算。由于空字符串 $\varnothing$ 使得补集运算产生了问题。正如前面所评论的那样，这种字符串可能在 0-型语言中，而不在像上面定义的语境敏感的语言、语境无关的语言或正则语言中。为了达到这一小节关于运算的目的，我们将把这些定义扩展到包括可能的空序列，并且，我们在这里不关心如何明确阐述生成这个空序列的产生式规则。给出这种扩展，我们以明显的方式定义补集。设 V_T 是一个有穷非空集，设 L 是一种短语结构语言，使得 $L \subseteq V_T^*$，即 L 是它的元素属于 V 的所有有穷序列的一个子集。因此，相对于 V_T，L 的补集是 $V_T^* - L$。

我们不加证明地陈述下列定理。

定理 4. 设 V_T 是一个有穷非空集。于是，V_T^* 的所有正则语言的子集族 L 形成了一个集合的布尔代数。

换言之，如果 L、$L' \subseteq V_T^*$，以及 L 和 L' 是正则语言，那么 $L \cup L'$、$L \cap L'$ 和 $V_T^* - L$ 也是正则语言。

这个定理对语境无关语言不成立，语境无关语言在交集或补集运算下是不闭合的。下面是关于交集的一个反例。设

$$V = V' = \{A, B, a, b, c\}$$

$$N = N' = \{A, B\}$$

$$P = \{A \to Ac, A \to B, B \to aBb, B \to ab\}$$

$$P' = \{A \to aA, A \to B, B \to bBc, B \to bc\}$$

$$S' = S。$$

于是，

$$L = L(G) = \{a^n b^n c^i \mid n \geqslant 1 \ \& \ i \geqslant 0\}$$

$$L' = L(G') = \{a^j b^n c^n \mid n \geqslant 1 \ \& \ j \geqslant 0\}$$

360 但是，

$$L(G) \cap L(G') = \{a^n b^n c^n \mid n \geqslant 1\}$$

而且，语言 $L(G) \cap L(G')$ 是非语境无关语言的一个标准例子，我这里没有证明这个事实(关于细节，参见 Aho and Ullman 1972, vol. 1, pp. 195 - 196)。语境无关语言在并集运算下是闭合的，我们将再回到这个事实。因此，从刚才给出的这个反例得出的结论是，这些语言在补集运算下不可能是闭合的，因为

$$L \cap L' = -(-L \cup -L')$$

我们把两种语言 L 和 L' 的积或串接定义为

$$LL' = \{x \mid x = uv, u \text{在} L \text{中}, v \text{在} L' \text{中}\}$$

因此，对于这四类语言，我们有下列定理。证明相当长，因此省略。

定理 5. 设 V_T 是一个有穷非空集。于是，作为 V_T^* 的子集的正则语言族、语境无关语言族、语境敏感语言族和 0 -型语言族，在并集和串接运算下，各自都是闭合的。

不可解问题。关于形式语言的一个问题被说成是不可解

的，当且仅当，根本没有确定这个问题的算法或机制决策程序。这类问题通常被说成在逻辑上是**不可判定的**。例如，在基本数论中任意给出的句子是否是可证明的，是不可判定的——哥德尔的著名结果之一。一阶逻辑的一个任意公式是否是有效的，更是无法确定的——丘奇和图灵各自独立得出的结果。下面选录关于语境无关语言的某些结果——至于更详细的讨论，参见霍普克罗夫特(Hopcroft)和厄尔曼(Ullman)的著作 (1969，第14 章)。

(i) 两种任意语境无关语言的交集是否是空的，是不可解的；

(ii) 一种语境无关文法在其终结词表上是否能生成所有字符串的集合，是不可解的；

(iii) 两种语境无关文法是否能生成相同的语言，是不可解的。

我们现在转向有歧义的某些结果。我们首先定义**最左派生**概念。我们说，一种派生是**最左的**，当且仅当，在每一步(n)，由这种文法的一个产生式规则替代的变量，在步骤(n)的字符串中，它的左边没有变量。例如，如果在一种派生中，(n)步有字符串 BA，用产生式规则 $A \rightarrow a$ 派生出($n+1$)步的字符串 Ba，那么，这种派生就不是最左的。能够表明，一种语境无关的文法中的一个终结字符串 u 的任意派生都能被 u 的最左派生所取代。一种语境无关的文法 G 被说成是**有歧义的**，当且仅当，在 $L(G)$ 中存在着一个字符串，这个字符串至少有两种不同的最左派生。一种语境无关的语言 L 被说成是**在本质上有歧义的**，当且仅当，生成 L 的每一种语境无关的文法都是**有歧义的**。于是，我们能证明关于歧义的定理。

定理 6. 任意一种语境无关的文法是否是有歧义的，是不可　361

解的。此外,一种语境无关的文法是否生成了一种本质上有歧义的语境无关语言,是不可解的。

这个定理的一个后果是,根本不可能自动地检查一种语境无关的新程序语言中是否有歧义。必须通过非算法的证明方法,才能确定没有歧义,或者,证明不是这种情况。

自然语言的应用。在乔姆斯基(1952)引入转换之后,人们认为多年来,根本没有一种短语结构文法能合理地处理像英语那样的自然语言的复杂性。另一方面,我们在定义真正有能力的转换时,带来了很多问题,因为早已表明,当把相当简单的转换增加到语境无关的语言中时,能够把非递归的语言转换成一种递归可枚举的语言,尽管一般认为,自然语言至少应该是递归的。

在过去的几十年内,关注的中心已经转回到了短语结构文法,例如,具有在布雷斯男(Bresnan,1982)和卡普兰(Kaplan)与布雷斯男(1982)的文章中提出的词汇功能文法形式的文法。在这条进路中,通过强迫补充文法结构的一种功能结构增加了由语境无关文法生成的句子结构。从形式的观点来看,我们能够把这种功能结构看成是与语境无关文法的派生树的相邻结点之间的附加约束一样强加的一个功能方程的集合。对于词汇功能文法接受的一个字符串来说,这些功能方程必须有惟一解。我们能够表明,一般说来,这种功能结构的强加意味着等价于一种语境敏感的文法。

在许多方面,广义的短语结构文法的发展是一个还要更有前途的方向。在加兹达尔(Gazdar)、克莱因、普卢姆(Pullum)和萨格(Sag)的著作(1985)中给出了初步系统的阐述。他们考虑的文法类弱等价于前面定义的语境无关的文法,但当然不是结构上等价。关于最近英语文法的一种概述,参见胡德莱斯顿

(Huddleston)和普卢姆的著作(2002)。

8.2 文法的表征定理

为了揭示证明正则文法的表征定理所需要的方法,我从有穷自动机及其接受的语言开始。

有穷自动机。我在本节将给出有穷自动机的几个定义和几个例子。因为一般的定义太简单和太直接,所以,我从定义出发,然后,举例说明。第一个定义接近于雷宾(Rabin)和斯科特的定义(1959)。

定义 1. 一个结构 $\mathfrak{A}=(A, V, M, s_0, F)$是一个有穷(确定型)自动机,当且仅当,

(i) A 是一个有穷非空集($\mathfrak{A}$ 的状态集),

(ii) V 是一个有穷非空集(字母表或词汇表),

(iii) M 是从笛卡儿积 $A \times V$ 到 A 的一个函数(M 定义了 $\mathfrak{A}$ 的转换表), 362

(iv) s_0 在 A 中(s_0 是 $\mathfrak{A}$ 的初始状态),

(v) F 是 A 的一个子集(F 是 $\mathfrak{A}$ 的最终状态集)。

这个定义的普遍性是显而易见的。从一个有穷自动机的一般概念的特定意义上说,这也是弱点的体现。惟一的限制是集合 A 和 V 是有穷的。另一方面,这些有穷的限制是关键的。

显然,我们能有定义 1 的某些极其平凡的模型。最简单的模型是:$A = \{s_0\}$, $V = \{0\}$, $M(s_0, 0) = s_0$, $F = \{\ \}$。

也许,下面是一个有穷自动机的最简单的非平凡事例。我们有两个字母的字母表和两种内部状态,我们也可以用符号“1”和“0”来表示这两个字母,用 s_0 和 s_1 表示这两种内部状态。自动机的转换函数根据下表定义为

	s_0	s_1
s_0 0	✓	
s_0 1		✓
s_1 0		✓
s_1 1	✓	

最后,我们选择内部状态 s_1 作为最终状态集 F 的惟一项。我已经说过,这是最简单的非平凡自动机。我的意思是说,转换表依赖于内部状态和输入的字母。站在更一般的概念立场上,这种方法本身显然是相当平凡的。

如果我们对很不平凡的、能起重要作用的、还是简单到容易直接描述的有穷自动机例证的日常经验作出考察,那么,很显然,最明显的应用领域之一是初等算术领域。直观上,我们大多数人都承认,我们在小学接受如何做算术运算的教育时,大致是有机制的或算法的。不需要任何形式的定义,通常很显然,我们能以一种完全客观的和机械的方式来定义算术的标准算法,并让一台很笨的机器进行运算。实际在直观意义上,这台机器很不机灵,因为当它面临相似的情境时,它总是做全然相同的事情,在它的结构中丝毫没有考虑到策略或意外。我这里的说法是含糊的,我只打算表达我们所有的人都很了解的一种直观思想。

当我们考虑通常如何做算术题时,定义 1 就有一种困难。我们得输出一个答案,而不只是移向最终状态。就理论工作和许多应用而言,定义 1 很简单。然而,寻求第二种密切相关的定义也是自然的。稍后我们将简要地研究这两种定义之间的关系。在这个新的定义中,我们排除了最终状态集,引入一个输出函数和输出词汇表或字母表。

363 **定义 2.** 一个结构 $\mathfrak{A}=(A, V_I, V_O, M, Q, s_0)$ 是一个带

输出的有穷(确定型)自动机,当且仅当,

(i) A 是一个有穷非空集,

(ii) V_I 和 V_O 都是有穷非空集(分别是输入词汇表和输出词汇表),

(iii) M 是从笛卡儿积 $A \times V_I$ 到 A 的一个函数(M 定义了转换表),

(iv) Q 是从笛卡儿积 $A \times V_I$ 到 V_O 的一个函数(Q 是输出函数),

(v) s_0 在 A 中(s_0 是初始状态)。

作为一个带输出的有穷自动机例子,即定义 2 意义上的一个有穷自动机,我们可以描述将在标准十进制中完成两位整数的按列相加运算的一个自动机。

$$A = \{0, 1\}$$

$$V_I = \{(m, n): 0 \leqslant m, n \leqslant 9\}$$

$$V_O = \{0, 1, \cdots, 9\}$$

$$M(k, (m, n)) = \begin{cases} 0 & \text{如果} \quad m + n + k \leqslant 9, \\ 1 & \text{如果} \quad m + n + k > 9, k = 0, 1 \end{cases}$$

$$Q(k, (m, n)) = (k + m + n) \bmod 10$$

$$s_0 = 0$$

因此,这个自动机先加第一列[①],如果没有"进位",内部状态存储为 0,如果有"进位",内部状态存储为 1,输出个位数的模的和 10,然后移动到两个十位数的输入,等等。初始内部状态 s_0 是 0,因为这道题开始时并没有"进位"。对完成其他算术运算的类

① 即个位数相加。——译者

似的自动机的描述,作为练习留给读者。

因为表面上定义 1 和定义 2 似乎是定义了稍微不同的方法,所以,有用的是,对定义 1 意义上的有穷自动机和带输出的有穷自动机(正如定义 2 中所定义的那样)之间的关系作出某种研究。

为了达到这个目的,将渴望引入这两种意义的有穷自动机的标准的同构概念。

定义 3. 设 $\mathfrak{A} = (A, V, M, s_0, F)$ 和 $\mathfrak{A}' = (A', V', M', s_0', F')$ 都是有穷自动机。于是,$\mathfrak{A}$ 和 $\mathfrak{A}'$ 是同构的,当且仅当,存在一个函数 f,使得

(i) f 是一一对应的,

(ii) f 的定义域是 $A \cup V$,值域是 $A' \cup V'$,

(iii) 对于 $A \cup V$ 中的每个 a,

$$a \in A,\text{当且仅当},f(a) \in A,$$

(iv) 对于 A 中的每个 s 和 V 中的每个 σ,

$$f(M(s, \sigma)) = M'(f(s), f(\sigma)),$$

(v) $f(s_0) = s'_0$,

364 (vi) 对于 A 中的每个 s,

$$s \in F,\text{当且仅当},f(s) \in F'。$$

显然,这个定义的条件(i)~(iii)意味着,对于 $A \cup V$ 中的每个 a,

$$a \in V,\text{当且仅当},f(a) \in V',$$

因而,关于 V 的这个条件没有必要陈述。从同构的一般代数概念或集合论概念的观点来看,更自然的是,根据一个基本集合 $B = A \cup V$ 定义一个自动机,然后要求,A 和 V 都是 B 的子集。

雷宾和斯科特(1959)通过不使 V 成为自动机的一部分回避了这个问题。他们相对于一个字母表 V 来定义一个自动机的概念 $\mathfrak{A}=(A, M, s_0, F)$,但就我们的目的而言,同样令人向往的是,在 $\mathfrak{A}$ 的定义中包括字母表 V,以便弄清楚在问题的一般框架中字母表的自然位置,而且,首先提供从一个字母表 V 到另一个字母表 V' 的一个简单方案。不管怎样,确切地说,如何解决这些问题不是最重要的。

定义 4. 设 $\mathfrak{A}=(A, V_I, V_O, M, Q, s_0)$ 和 $\mathfrak{A}'=(A', V'_I, V'_O, M', Q', s'_0)$ 是带输出的有穷自动机。于是,$\mathfrak{A}$ 和 $\mathfrak{A}'$ 是同构的,当且仅当,存在一个函数,使得

(i) f 是一一对应的,

(ii) f 的定义域是 $A \cup V_I \cup V_O$,值域是 $A' \cup V'_I \cup V'_O$,

(iii) 对于 $A \cup V_I \cup V_O$ 中的每个 s,

$$s \in A,\text{当且仅当},f(s) \in A',$$

(iv) 对于 $A \cup V_I \cup V_O$ 中的每个 σ,

$$\sigma \in V_I,\text{当且仅当},f(\sigma) \in V'_I,$$

(v) 对于 A 中的每个 s 和 V_I 中的每个 σ,

$$f(M(s, \sigma)) = M'(f(s), f(\sigma)),$$

(vi) 对于 A 中的每个 s 和 V_I 中的每个 σ,

$$f(Q(s, \sigma)) = Q'(f(s), f(\sigma)),$$

(vii) $f(s_0) = s'_0$。

定义 3 和定义 4 是如此相似,把它们都明确地写出来,是有点多余的。然而,根据摆在我们面前的这两个定义,应该显而易见的是,定义相似结构的同构的一般集合论方法,在这里也是适用的。在通常意义上,这两个同构定义是与公理无关的,它们不

依赖于定义 1 或定义 2 中的根本假设。正如已经评论的那样，这些假设只是有穷性假设。因为这个原因，在这两个例子中，这两个同构定义的与公理无关的特征并不很明显。

有穷自动机接受的语言。我们希望描述一个给定的有穷自动机能够接受或识别的语言。这时给出的定义，在定义 1 的意义上，是针对一个有穷自动机 $\mathfrak{A} = (A, V, M, s_0, F)$ 的定义。

365 首先，V^* 是 V 的元素的所有有穷序列的集合，包括空序列 $\varnothing$ 在内。V^* 的元素被称为句子、字符串或纸带。如果 $\sigma_1, \cdots, \sigma_n$ 在 V 中，那么，$\sigma_1\sigma_2\cdots\sigma_n$ 在 V^* 中，换言之，我们经常通过并列 V 的元素名表明 V^* 的元素。

其次，根据 A 中的 s、V^* 中的 x 和 V 中的 σ 的下列递归定义，我们能够把函数 M 推广到从 $A \times V^*$ 到 A 的一个函数，

$$M(s, \varnothing) = s$$

$$M(s, x\sigma) = M(M(s, x), \sigma)$$

这样，得到两种状态，即两个字母的字母表自动机

	s_0	s_1
s_0 0	√	
s_0 1		√
s_1 0		√
s_1 1	√	

让我们运用这个递归定义针对字符串 $x = 101$ 计算 M。

$$\begin{aligned} M(s_0, 101) &= M(M(s_0, 10), 1) \\ &= M(M(M(s_0, 1), 0), 1) \\ &= M(M(s_1, 0), 1) \\ &= M(s_1, 1) \\ &= s_0 \end{aligned}$$

因此,字符串 101 没有被接受,因为最终状态是 s_0,而不是 s_1。在更形式的意义上,我们有:

定义 5. $\mathfrak{A}$ 接受 V^* 的一个字符串 x,当且仅当,$M(s_0, x)$ 在 F 中。

我们通常将把 $\mathfrak{A}$ 接受的字符串称为 $\mathfrak{A}$ 的**句子**。$\mathfrak{A}$ 接受的语言是 $\mathfrak{A}$ 的所有句子的集合 $L(\mathfrak{A})$。在文献中,$\mathfrak{A}$ 接受的语言通常被称为 $\mathfrak{A}$ **接受的纸带的集合**(the set of tapes)。

存在着一个自动机的自然的等价概念,这个概念比上面定义的严格同构概念更弱一些。

定义 6. 两个自动机(在弱的意义上)是等价的,当且仅当,它们接受同样的语言。

说自动机的同构关系总是比等价关系更强是错误的,因为两个自动机显然能是同构的,而不能是等价的,由于词汇表或字母表不同。我们能够轻易地改变定义 6 来说明这个事实。我们不要求语言相同,只要求从 V 到 V' 的函数是一一对应的。

一个有穷自动机是**连通的**,当且仅当,对于每个状态 s,都有一个字符串 x,使得 $M(s_0, x) = s$。于是,我们有明显的定理。

定理 1. 任意有穷自动机都等价于一个连通的有穷自动机。

下面的定理陈述了一个非空有穷自动机接受的语言的一个简单的有穷论标准。

定理 2. 一个有 n 个状态的有穷自动机接受的句子集是非空的,当且仅当,这个有穷自动机接受长度小于 n 的句子。 366

证明:如果接受了长度小于 n 的一个词,那么这种语言就是非空的,根据反证法的方式,设被接受的最短的句子 w 的长度大于等于 n。首先应该注意的一点是,在处理 w 时,$\mathfrak{A}$ 必须至少两次经过某个状态——这是立刻从任意字符串的 M 的定义得

出的结论。设这个状态是 s,我们能把 w 分解为三个部分：

$$w = w_1 w_2 w_3$$

并且 $w_2 \neq \varnothing$, 根据关于 s 的假设

$$M(s_0, w_1) = M(s, w_2) = s$$

和

$$M(s, w_3) \in F$$

现在,w_1 或 w_3 都不是空集,否则, $w_2 = w$, 并且,严格说来,我不根据 s 的重复,来分解 w 。因此,我们现在有

$$M(s_0, w_1 w_3) = M(s_0, w_1 w_2 w_3) \in F$$

和

$$| w_1 w_3 | < | w_1 w_2 w_3 |$$

与假设相反。

证毕。

更令人惊讶的是,当一个有穷自动机接受的语言是无穷的时,也有一个严格的有穷论的检验。

定理 3. 一个有 n 个状态的有穷自动机接受的语言(即句子集)是无穷的,当且仅当,这个自动机接受长度等于或大于 n 和小于 $2n$(即, $n < | w | < 2n$)的句子 w。

证明: 首先,假设存在一个被接受的语句 w, 并且, $n \leqslant | w | \leqslant 2n$。于是,正如定理 2 的证明一样,我们能够找到一个重复状态 s,使得

$$w = w_1 w_2 w_3$$

$$w_2 \neq \varnothing$$

$$M(s_0, w_1) = M(s, w_2) = s$$

$$M(s, w_3) = F,$$

然后得出，对于所有 i，

$$w_1 w_2^i w_3$$

被接受，这里，w^i 是长度为 i 的 w 的一个字符串，而且，句子的这个集合显然是无穷的。

另一方面，让我们现在假设，$\mathfrak{A}$ 接受一个无穷大的句子，但没有一个句子的长度在 n 和 $2n-1$(包括 $2n-1$)之间。设 w 是最小长度大于等于 $2n$ 的被接受的一个句子。于是，根据证明定理 2 时给出的论证，我们能够找到一个重复状态 s 和一个长度为 $1 \leqslant |w_2| < n$ 的句子 w_2，使得

$$w = w_1 w_2 w_3$$

并且接受 $w_1 w_3$。于是，要么 $n \leqslant |w_1 w_3| < 2n$，要么如果 $|w_1 w_3| \leqslant 2n$，能重复这个论证。在有限的步骤内，我们将获得 $w'_1 w'_3$，使得 $n \leqslant |w'_1 w'_3| < 2n$，与我们的假设相反。 367

证毕。

从定理 2 和定理 3 得出的结论是，存在着下列算法。

定理 4. *有一种算法来确定，一个有穷自动机接受的语言是空的、有穷但非空的，或无穷的。*

正则文法和有穷自动机。为了确立正则文法和有穷自动机之间的等价性，在许多方面自然引入了概率或非确定型自动机。我们前面没有考虑非确定型自动机。大致说来，我们能够通过忽略实际转换概率和惟一允许正向概率的跃迁，从一个概率自动机获得一个非确定型自动机。依我看来，这个概念很不直观，因此，我们应该使用概率自动机，而不是非确定型自动机，但我注意到，关于这一点，我这里采取的进路不同于标准文献中的许多进路。概率自动机的定义是对确定型自动机定义的一种直接概括。

定义 7. 一个结构 $\mathfrak{A}=\{A, V, p, s_0, F\}$ 是一个(有穷)概率自动机,当且仅当,

(i) A 是一个有穷非空集,

(ii) V 是一个有穷非空集,

(iii) p 是 $A\times V$ 上的一个函数,使得对于 A 中的每个 s 和 V 中的每个 σ,$p_{s,\sigma}$是 A 上的概率密度,即

(a) 对于 A 中的每个 s', $p_{s,\sigma}(s')\geqslant 0$,

(b) $\sum_{s'\in A} p_{s,\sigma}(s')=1$,

(iv) s_0 在 A 中,

(v) F 是 A 的一个子集。

设 $\mathfrak{A}=(A, V, p, s_0, F)$ 是刚才定义的一个概率自动机。于是,我们可以把 p 扩展到 V^*

$$p_{s,\varnothing}(s')=\begin{cases}1 & \text{如果 } s'=s\\ 0 & \text{其他}\end{cases}$$

$$p_{s,x\sigma}(s')=\sum_{s''\in A} p_{s,x}(s'')\cdot p_{s'',\sigma}(s')$$

如果 $\sum_{s\in F} p_{s_0,x}(s)>0$,即如果进入一个最终状态(即 F 中的一个状态)的概率是正的,那么 $\mathfrak{A}$ 接受了一个句子 x。注意,所有的变量都一定在接受的这个定义中,因此,我们可以把接受 V^* 中的字符串 x 的概率简单地写为 $p(x)$。运用这个接受定义,我们希望表明,一个概率自动机接受的语言也是一个有穷自动机接受的语言。稍后我们将简要地讨论一个不同的接受定义,对于这种定义来说,情况并非如此。

368 **定理 5.** 如果一种语言 L 被一个概率自动机所接受,那么它就被一个有穷自动机所接受。

证明: 最重要的一步是把有穷自动机的状态看成是概率自

动机的状态集。其余的部分是很容易的。

设 $\mathfrak{A}=(A, V, p, s_0, F)$ 是接受语言 $L(\mathfrak{A})$的一个概率自动机。

确定型自动机的定义如下：

$$A' = 2^A \qquad = A \text{ 的幂集},$$

$$V' = V \qquad (\text{相同的词汇表}),$$

$$s'_0 = \{s_0\},$$

$$M(B, \sigma) = \{s': p_{s,\sigma}(s') > 0 \text{ 对于 } B \text{ 中的某个 } s\},$$

对于 A'中的 B(即 B 是 A 的一个子集)和 V 中的σ,

$$F' = \{B: B \in A' \& B \cap F \neq \varnothing\}。$$

我们现在需要表明，$\mathfrak{A}' = (A', V', M, s'_0, F')$ 接受的语言与 $\mathfrak{A}$ 接受的语言相同。

我们通过对 x 长度的归纳，表明

$$M(s'_0, x) = B, \text{ 当且仅当},$$

$$p_{s_0, x}(s') > 0 \leftrightarrow s' \in B, \text{ 即}$$

$$M(s'_0, x) = \{s': P_{s_0, x}f(s') > 0\}。$$

如果 $|x| = 0$

$$M(s'_0, \varnothing) = s'_0 = \{s_0\}$$

和

$$p_{s_0, \varnothing}(s_0) = 1, \text{ 因此，没有别的证明。}$$

我们现在用关于 x 长度的强归纳。假设当 $|x| < n$ 时这个结果成立。然后考虑 $x\sigma$ 。通过归纳假设

$$M(s'_0, x) = \{s'': p_{s_0}, x(s'') > 0\} = B$$

根据 M 的定义

$$M(B, \sigma) = \{s' : p_{s'', \sigma}(s') > 0, \text{对于 } B \text{ 中的某个 } s''\}。$$

因此

$$M(s'_0, x\sigma) = \{s' : \sum_{s'' \in A} p_{s_0, x}(s'') p_{s'', \sigma}(s') > 0\},$$

即

$$M(s'_0, x\sigma) = \{s' : p_{s_0, x\sigma}(s') > 0\}。$$

最后,不难证明

$$M(s'_0, x) \in F, \text{当且仅当}, \sum_{s \in F} p_{s_0, x}(s) > 0,$$

因此

$$L(\mathfrak{A}) = L(\mathfrak{A}')。$$

证毕

然而,正如我们已经注意到的那样,我们可以强化概率自动369机的接受定义,得到一个不同的结果。设 $\mathfrak{A}$ 是一个概率自动机,设 $\lambda \in [0, 1]$。于是,$L(\mathfrak{A}, \lambda)$ 是含有割点 λ 的公认的句子集。在形式上

$$L(\mathfrak{A}, \lambda) = \{x : x \in V^* \,\&\, p(x) > \lambda\}。$$

雷宾的著作(1963)中有一个例子,就这个例子而言,$L(\mathfrak{A}, \lambda)$ 不是一种正则语言。如果 λ 是有理数,则 $L(\mathfrak{A}, \lambda)$是正则的。

现在直接进入主题。我们首先证明,某个有穷自动机接受任何一种正则语言。通过首先构造一个适当的概率自动机,然后运用前面的定理进行这个证明。自然对应的是,把非终结词汇表 N 映射到自动机的状态和把终结词汇表 V_T 映射到自动机的词汇表 V。大多数文法中都预计到非确定的或概率的跃迁,例如,我们有关于名词短语或动词短语的若干重写或产生式

规则。

定理 6. **如果 G 是一种正则文法，那么存在一个将接受 $L(G)$ 的有穷自动机。**

证明： 设 $G=(V,\ N,\ P,\ S)$ 是一种正则文法。对于这个自动机，

$$A = N \cup \{q\},\ q \notin N$$

(换言之，这些状态是 N 和一个其他符号 q)。

$$V' = V_T$$

$$s_0 = S$$

$$F = \{q\}。$$

我们现在把 p 定义为

(i) 如果 $\alpha \to a$ 是一个产生式，即在 P 中，则

$$p_{\alpha,a}(q) > 0,$$

(ii) 如果 $\alpha \to a\beta$ 在 P 中 $(\alpha,\ \beta \in N)$，则

$$p_{\alpha,\ a}(\beta) > 0,$$

(iii) $p_{\alpha,a}(q) = 1$。

因此 $\mathfrak{A}=(A,\ V',\ p,\ s_0,\ F)$ 是根据文法 G 构造的概率自动机。我们首先证明 $L(G) \subseteq L(\mathfrak{A})$。设 x 是 $L(G)$ 中的一个句子。于是，在这种形式的 G 中，有 $x = a_1 \cdots a_n$ 的一个派生：

$$S \Rightarrow a_1\alpha_1 \Rightarrow a_1a_2\alpha_3 \Rightarrow \cdots \Rightarrow a_1 \cdots a_{n-1}\alpha_{n-1} \Rightarrow a_1 \cdots a_n$$

从作为(i)—(iii)给出的离散概率分布 p 的定义来看，显然，这种派生的每一步都有正向概率，比如说，$\varepsilon_i > 0$，因此

$$\prod_{i=1}^{n} \varepsilon_i > 0,$$

据此，根据概率自动机的接受定义，$L(\mathfrak{A})$接受 x。

我们下一步根据类似的论证表明，$L(\mathfrak{A}) \subseteq L(G)$。设 $x \in L(\mathfrak{A})$。于是，存在状态 S 的一个序列，$\alpha_1, \cdots, \alpha_{n-1}, q$，使得 $p_{S, a_1}(\alpha_1) > 0$，$p_{\alpha_{i-1}}, \alpha_i(\alpha_i) > 0$，$p_{\alpha_{n-1}, a_n}(q) > 0$，当然，这里

370 $x = a_1 \cdots a_n$。但从 p 的定义立刻得出的结论是，在 G 中一定有 x 的派生，因此，$x \in L(G)$。

给定所进行的证明，也就是说，如果 G 是一种正则文法，那么存在着一个将接受 $L(G)$的概率（有穷）自动机，我们运用定理 5 断定，$L(G)$被某个有穷自动机所接受。

证毕。

定理 7. 已知一个有穷自动机 $\mathfrak{A}$，存在一种正则文法 G，使得 $L(G) = L(\mathfrak{A})$。

证明：设 $\mathfrak{A} = (A, V, M, s_0, F)$ 是一个有穷自动机。于是，我们把一种文法 $G = (V', N, P, S)$ 定义为

$$V' = A \cup V$$

$$N = A \qquad (\text{So } V'_T = V)$$

$$S = s_0。$$

产生式规则的集合 P 是关键项，但是，从逆定理（定理 6）的证明来看，这种定义的方式应该是清晰的：

(i) 如果 $s, s' \in A$, $a \in V$, $M(s, a) = s'$，则 $s \to as' \in P$；

(ii) 如果 $s \in A$, $a \in V$，对于某个 $s' \in F$，$M(s, a) = s'$，则 $s \to a \in P$。

当然，只有形式(i)或(ii)的产生式能够在 P 中。

我们留作一种练习，让读者来完成这个证明。如前所述，很自然的是，把其余要证明的部分分成两部分：

证明 $L(\mathfrak{A}) \subseteq L(G)$ 和证明 $L(G) \subseteq L(\mathfrak{A})$。

证毕。

定理 6 和定理 7 彼此根据对方共同提供了正则文法和有穷自动机的一种相互表征。原则上,由此为正则文法提供了一种程序说明,但是,仔细看看这两个定理的证明表明,这些表征结构太接近于最初的结构,不可能给出很丰富的表征印象。令人怀疑的是,任何一个人都会看着这种表征,然后说,"啊,我现在明白了,一种正则文法是如何起作用的。"一个例外是,这个表征定理确实表明,一个严格的有穷自动机能够处理一种正则文法,而且,从这些文法形式来看,这未必是显而易见的。

许多人可能觉得,归功于克伦(Kleene, 1956)的关于正则语言的下列结构定理,提供了一种更有趣的表征,即使它在特征上不是程序性的。下一节试图进行一种更加心理学的处理分析。

定理 8. (Kleene,1956)设 V_T 是一个有穷非空集。于是,正则语言或有穷自动机接受的等价语言 V_T^* 的子集类,是包括 V_T^* 的所有有穷子集和在并集、串接和闭合条件下闭合的最小类。

一种语言 L 的闭包是由 L 中的串接字符串产生的所有有穷字符串的集合。例如,如果 $L = \{1, 01\}$,那么 L 的闭包是 $\{1, 01, 11, 011, 101, 111, \cdots\}$。在这个闭包定义中通常包括空字符串在内。关于这一点参见下一段。串接的定义刚才在定理 5 之前就给出了。 371

关于空序列的评论。在正则文法、语境无关文法和语境敏感文法的定义中,我们不允许下列形式的产生式规则:

$$\alpha \to \varnothing$$

这里,$\alpha \in V^*$。人们有可能把问题扩展到包括这个特殊规

则——或者说,这种特殊形式的规则。在这些定理的证明中,比如,刚才考虑的正则语言的两个定理的证明,必须对它们的出现或缺少作出冗长的解释。关于空序列类的产生式规则的好的讨论,参见霍普克罗夫特和厄尔曼的著作(1969, pp. 15 ff.)。

下推自动机和语境无关的语言。首先,我们定义一个概率下推自动机。概率部分是对概率有穷自动机概念的一种自然概括。这个新思想和重要思想是下推存储的思想,这是一种受限制的记忆形式。这个自动机能够在这个储存器中存储结构信息,但是,它会使状态只根据存储器的顶符号进行状态转换——存储器的运算就像自助餐厅里的一摞盘子一样根据"先进后出"的原则进行。这个机器配备有一个特殊的存储词汇表。转换函数现在不仅依赖于输入符号和当前状态,而且依赖于存储器的顶符号。当然,在概率情况下,我们拥有依赖于下列三个条件的一个条件概率分布:s 在 A 中、α 在 V 中和 γ 在 Γ(存储词汇表)中。

通过忽略概率分布,从一个概率自动机能获得一个不确定或非确定下推自动机。

定义 8. 一个结构 $\mathfrak{A} = (A, V, \Gamma, p, s_0, \gamma_0, F)$ 是一个概率下推自动机,当且仅当,

(1) A 是一个非空有穷集(状态集)。

(2) V 是一个非空有穷集(输入词汇表)。

(3) Γ 是一个非空有穷集(存储词汇表)。

(4) 对于 A 中的每个 s、$V \cup \{\varnothing\}$ 中的每个 σ 和 Γ 中的每个 γ,$p_{s,\sigma,\gamma}$ 是 $A \times \Gamma^*$ (概率转换表)的一个有穷子集上的一个概率密度。

(5) $s_0 \in A$ (初始状态)。

(6) $\gamma_0 \in \Gamma$ (起始存储符)。

(7) $F \subseteq A$ (最终状态集)。

关于定义 8 的几个评论是有顺序的。首先，尽管在 p 的定义中包含了 Γ^*，不只是 Γ，但是，只能存储确定的有限数目的不同词（Γ 的元素），即使它们的序列可能是无限长的。第二，没有输入，状态转换也能发生，因为 σ 可能是 $\varnothing$，正如从这个定义的第四条能看到的那样。第三，我们通过用从 $A \times (V \cup \{\varnothing\}) \times \Gamma$ 到 $A \times \Gamma^*$ 的一种关系替代条件概率密度 p，能从一个概率式下推自动机得到一个非确定自动机： 372

$$s, \sigma, \gamma M s', \alpha, \text{当且仅当}, p_{s,\sigma,\gamma}(s', \alpha) > 0,$$

其中，$\alpha \in \Gamma^*$。[从(4)得出，M 的值域是有穷的。]

一个对 (s, α)，其中 s 在 A 中，α 在 Γ^* 中，是一个确定的构型。我们说，如果 $\mathfrak{A}$ 在状态 s 中，α 是下推存储器中的字符串，α 最左边的符号是存储器的顶符号，那么一个下推自动机 $\mathfrak{A}$ 在构型 (s, α) 中。我们用符号

$$\sigma, (s, \gamma\alpha) \vdash (s', \beta\alpha)$$

意指，给定 V 中的输入 σ 和构型 $(s, \gamma\alpha)$，自动机转到构型 $(s', \beta\alpha)$，其中，$\beta \in \Gamma^*$。（这里“转到”的概率是正的。）

我们以相同的方式把这个定义扩展到 V^* 中的输入序列 x，正如我们扩展有穷自动机的转换函数那样。特别是，我们通过归纳定义：

$$\varnothing, (s, \gamma) \vdash^* (s, \gamma)$$

和

$$x\sigma, (s, \gamma) \vdash^* (s', \alpha'), \text{当且仅当}, \exists s'' \text{在} A \text{中}, \alpha'' \text{在} \Gamma^* \text{中}$$

使得

$$x, (s, \gamma) \vdash^* (s'', \alpha'') \text{和} \sigma, (s'', \alpha'') \vdash (s', \alpha')。$$

在技术上，这是关于这个定义的一个明显的观点。我们希

望允许 x 的项是空位或$\varnothing$。因而实际上,x 在$(V^* \cup \{\varnothing\})$中。我们把细节留给读者。

对于一个下推自动机 $\mathfrak{A}$ 来说,我们把**最终状态接受的语言**定义为

$$L(\mathfrak{A})=\{x: x, (s_0, \gamma_0)\vdash^* (s, \alpha), \text{对于 } s \text{ 在 } F \text{ 中和 } \alpha \text{ 在 } \Gamma^* \text{ 中}\}$$

一个直观上吸引人的密切相关的定义是**被空存储接受的语言**:

$$L_E(\mathfrak{A}) = \{x: x, (s_0, \gamma_0) \vdash^* (s, \varnothing), \text{对于 } A \text{ 中的任意 } s\}$$

我们需要一个初步结果。

定理 9. *一种语言被某个下推自动机通过最终状态所接受,当且仅当,它被某个下推自动机通过空存储器所接受。*

鉴于定理 9,我们可以毫不含糊地言说被下推自动机接受的一种语言,不用通过最终状态或存储器来指定说明。

定理 10. *如果一种语言是语境无关的,那么存在一个接受这种语言的下推自动机。*

证明: 由于定理 3,在不失一般性的前提下,我们可以把这种语言的文法 $G = (V, N, P, S)$ 看成是格雷巴赫范式。构造处理 $L(G)$的下推自动机 $\mathfrak{A}$ 的最启迪人的方面是,$\mathfrak{A}$ 只需要有一种状态:

$A = \{s_0\}$, $V' = V_T$, $\Gamma = N$, $\gamma_0 = S$, $F = \varnothing$(因此我们通过空存储接受)以及

$$p_{s_0, s, \gamma}(s_0, a) > 0, \text{当且仅当}, \gamma \to sa \in P。$$

373 通过表明下面一点的归纳论证完成这个证明,

$$S \Rightarrow_G^* x, \text{当且仅当}, x, (s_0, \gamma_0) \vdash^* (s_0, \varnothing)。$$

证毕。

我们把这种论证留作练习。

正如这个证明所表明的那样,下推存储实际上正在做所有计算的工作,所以,事实上,语境无关的语言能够通过一个窄类的下推自动机,即只有一种状态的那些自动机来表征。

定理 11. 如果 $\mathfrak{A}$ 是一个下推自动机,那么 $L(\mathfrak{A})$ 是一种语境无关的语言。

证明: 设 $\mathfrak{A} = (A, V, \Gamma, p, s_0, \gamma_0, \varnothing)$,因此,$\mathfrak{A}$ 通过空存储器接受。我们把接受 $L(\mathfrak{A})$ 的语境无关文法定义为

$$V'_T = V$$

$$N = A \times \Gamma \times A \cup \{S\},\ S \in A \times \Gamma \times A。$$

因此,$V' = N \cup V'_T$。

更重要的是,注意,除了 S 之外,N 的元素是三元组 。P 中的产生式规则有三种形式:

(i) 对于 A 中的每个 s,$S \to (s_0, \gamma_0, s)$;

(ii) $(s, \gamma, s') \to \sigma$,当且仅当,$p_{s, \sigma, \gamma}(s', \varnothing) > 0$;

(iii) $(s, \gamma, s') \to \sigma(s_1, \gamma_1, s_2)(s_2, \gamma_2, s_3)\cdots(s_m, \gamma_m, s')$;当且仅当,$p_{s, \sigma, \gamma}(s_1, \gamma_1\cdots\gamma_m) > 0$,其中,$\sigma \in V \cup \{\varnothing\}$。

通过表明下面这点的归纳论证来完成证明

$S \Rightarrow_G^* x$,当且仅当,对于 A 中的某个 S,x,$(s_0, \gamma_0) \vdash^* (s, \varnothing)$。

——在形式上与定理 10 的证明非常类似,但稍微更复杂些。我们把这一步留作练习。

证毕。

图灵机和线性有界自动机。首先,我们需要引入图灵机。[①] 一个图灵机有一个读写头。给定一个内部状态和一个输入字母

---

① 在与寄存器机和可计算性有关的 3.5 节,简要地讨论过图灵机。

的扫描,图灵机通过下列方式执行一次移动:

(i) 改变内部状态,

(ii) 在被扫描的单元上打印一个非空位符号,从而代替被扫描的符号,

(iii) 把它的磁头向左或向右移动一个单元。

定义 9. 一个结构 $\mathfrak{A} = (A, V, \Gamma, M, s_0, F)$ 是一个图灵机,当且仅当,

(1) A 是非空的和有穷的(状态集),

(2) V 是非空的和有穷的(输入字母表),

(3) Γ 是非空的和有穷的,并且,$V \subseteq \Gamma$(带有空位符号 B 的纸带符号集在 $\Gamma - V$ 中),

(4) $M: A \times \Gamma \rightarrow A \times (\Gamma - \{B\}) \times \{L, R\}$,这里 M 是下一个移动函数,

(5) $s_0 \in A$(起始状态),

(6) $F \subseteq A$(最终状态集)。

374 **一个线性有界自动机**是保存在放置输入的纸带小格子内的一个图灵机。在接下来的两个定理中给出了语境敏感的语言和0-型语言的表征定理,我们不证明这两个定理。由于各种原因,已经讨论过的正则语言或语境无关语言的表征,在概念上有更大的重要性。另一方面,图灵机和部分递归函数之间的相互表征具有很大的重要性。第3章已经证实了这一点。

定理 12. 一种语言是语境敏感的,当且仅当,它被某个线性有界自动机所接受。

定理 13. 一种语言是0-型的,当且仅当,它被某个图灵机所接受。

最后,我们陈述关于语境敏感语言的一个重要定理。我们说,一种文法 $G = (V, N, P, S)$ 是递归的,当且仅当,有一种

算法来决定，V_T^* 中的每个字符串 x 是否是 $x \in L(G)$。

定理 14. *任意语境敏感的文法都是递归的。*

证明的基本思路：因为词汇表 V 是有穷的，所以，产生式规则的总数是有穷的，并且，一种语境敏感的文法的产生式规则绝不会减少字符串的长度，因为对于一个长度为 n 的字符串 x 来说，很容易算出一个上界，作为关于能够派生长度为 n 的字符串个数 n 的一个函数。（如果像在 0-型文法中那样，允许删除，就不可能是这种情况。）

8.3 有穷自动机的刺激-反应表征

本节的中心思想十分简单——将表明，根据应用公认的条件作用原理，通过一个适当的强化程序表，如何能在理论上训练一个生物体，使他作出像一个有穷自动机那样的反应。为了表明能够使遵守刺激条件作用和抽样的一般定律的一个生物体适应成为一个自动机，首先需要在常用的心理学概念内解释一个字母表中的一个字母的概念和内部状态概念。在我自己关于这些问题的思考中，我最初也许被自然的企图所误导：认为自动机的内部状态等同于生物体的条件作用状态。然而，这种思想原来证明是明显错误的。首先，生物体的各种可能的条件作用状态对应于能够使生物体适应成为的各种可能的自动机。大体上说，每个条件作用状态都对应于一个不同的自动机。大概接下来最自然的想法是考虑一个特定的条件作用状态，并用个别刺激的条件作用表示这个自动机的内部状态。在很受限制的情况下，这种对应是可行的，但在一般情况下却不可行，其原因下面会明白。证明可行的对应关系是：自动机的内部状态等同于

生物体的反应。

在自动机将要接受的字母表中的字母和在刺激-反应理论中适当的对象之间的这种对应,是相当明显的。字母表中的字母以一种自然的方式对应于在一次特定试验中为一个生物体呈现的刺激元素的集合。这看起来像是幸运的,但是,自动机的输入和呈现给生物体的刺激之间的对应,以及计算机的内部状态和生物体的反应之间的对应,在概念上是很自然的。

375

因为语言学家为了促进心理学理论的未来发展所提出的那些问题具有概念上的重要性,或许,首先因为语言行为是我们行为模式的最有人性的特征,所以,重要的是使关于下列刺激-反应理论能够作出的主张尽可能明确,即为了说明包括语言行为在内的复杂行为,这个理论的基本概念似乎很简单,但对于许多人而言,似乎很不充分。我举两个例子,这两个例子体现了很有用的类比。第一个例子是,把标准的数学简化为集合概念和是一个集合项的元素的一种简单关系。从朴素的观点来看,能把复杂的高等数学简化为与集合隶属关系一样简单的关系,似乎是不可信的。但无疑确实如此,我们在细节上知道如何能够作出这种简化。这不是建议说,比如,一位数学家在思考一个数学问题乃至在明确地表述和证明它时,只根据关于集合隶属的永无止境的复杂陈述进行运算。通过适当明确的定义,我们引入了许多附加概念,即在论述中实际用到的概念。然而,事实仍然是,能够简化为单一的集合隶属关系,而且,这种简化实际在细节上已经实现了。第二个例子是计算机的简单的机器语言的状况,这个例子接近于我们当前的研究。再一次从朴素的观点来看,似乎难以令人置信的是,当现代计算机的基本语言本质上仅仅是由 1 和 0 的有穷序列组成时,计算机能根据信息处理或数字计算来做它们能做的事情;但已经引入的更复杂的计算机语

言根本不是为了机器的方便,而是为了人类使用者的方便。任何一种更复杂的语言,诸如 Visual C 语言,如何能通过编程序或其他方法简化为一种简单的机器语言,是很清楚的。似乎对我来说,同样的态度也适合于刺激-反应理论。我们无法指望在刺激-反应关系中直接处理复杂的人类行为。正如在刚才提到的两个例子中那样,我们可以希望构造一个令人满意的系统理论,据此,能够引入新的明确定义的一个链条乃至更复杂的概念。正是这些新的明确定义的概念将与更复杂的行为方式直接相关。

在转向特殊的数学进展之前,有用的是搞清楚,如何能用本节的发展表明许多普通的条件作用概念,特别是条件作用只指像眨眼之类的简单反应的主张是错误的。这种错误是把这个基础理论限于的特殊应用误认为是该理论本身的范围。从更广泛的概念立场来看,关于经典条件作用的实验确实代表了实验的一个很小的范围。然而,重要的是意识到,关于经典条件作用的实验不能定义条件作用理论本身的范围和界线。本节的主要目的是表明,任意有穷自动机,不管多么复杂,都可以完全在刺激-反应理论的范围内被构造出来。但从自动机的观点来看,经典条件作用代表了一个特别平凡的自动机的例子。经典条件作用可以通过有一个字母的字母表和一种内部状态的自动机来表示。下一个最简单的情况对应于经典鉴别实验的结构。这里,字母表中不止有一个字母,但自动机的转换表不可能依赖于自动机的内部状态。在鉴别的情形中,我们可以再一次把反应看成对应于自动机的内部状态。在这种意义上,与经典条件作用的情形相反,不止有一种内部状态,但最基本的是,自动机的转换表不依赖于内部状态,而只依赖于根据试验者确定好的程序呈现出的外部刺激。最重要的是意识到,正如在经典条件作用 376

实验的情况下那样，这种限制不是在任何意义上都内在于条件作用理论本身的一种限制。它只代表了对所限于的某一实验类的关注。

即使技术细节留在后面来谈，还是有可能举出一个很明确的条件作用的例子，这个例子超越了经典的情形，但也许代表了最简单的非平凡自动机。我说的非平凡的意思是指：在这个字母表中不止有一个字母；不止有一种内部状态；并且自动机的转换表是外部刺激和当前内部状态的一个函数。我们可以以正在一个迷宫里跑的一只老鼠为例。设计对老鼠的强化程序表，使得老鼠成为一个两种状态的自动机。我们用由一张黑卡或一张白卡组成的两个字母的一个字母表作为自动机的外部字母表。这个迷宫的每一个选择地点将由左转或右转构成。在每个选择地点，要么呈现一张黑卡，要么呈现一张白卡。下表描述了自动机的强化程序表和转换表。

	L	R
LB	1	0
LW	0	1
RB	0	1
RW	1	0

这样，第一行表明，当前一种反应是转向左(L)和在选择地点呈现出一张黑色刺激卡(B)时，强化老鼠以概率1左转。第二行表示，当前一种反应是左转和在选择地点呈现出一张白色刺激卡时，这时强化老鼠100%右转，另外两种可能，依此类推。从形式的观点来看，这是一个简单的强化程序表，但是，关于前一种反应和呈现的刺激卡已有的偶然性的两个方面，使得这个程序表在许多方面比平常跑的老鼠的强化程序表更复杂。

没有借口说,这个简单的两种状态的自动机在任何一种意义上都能胜任认真的学习。例如,我不是在提议说,有许多机会把一种很简单的正则文法教会老鼠。而我是在向心理学家建议说,从每种动物能做什么的角度来看,状态数不多的自动机已经面临着直接的实验挑战。例如,可以训练一只猴子模仿的最复杂的自动机是什么呢?在这种情况下,似乎有某种可能至少接近相当复杂的正则文法(运用上一节的定理 10,即通过一个有穷状态的自动机,任意正则文法都是可表示的)。在低等动物情况下,为了使它们能模拟的自动机的复杂性最大化,有必要探索生物体最敏感和最易作出反应的最全面的刺激类型。 377

刺激-反应理论。这里给出的刺激反应理论的表述密切遵循了苏佩斯的文章(1969b)。这个理论基于六个原始概念,每个原始概念都有一个直接的心理学解释。第一个原始概念是刺激的集合 S,我们将假定这个集合是非空的,但是,我们在一切场合都不限制它要么是有穷集,要么是无穷集。第二个原始概念是反应的集合 R,第三个原始概念是可能强化的集合 E。正如在刺激集合情形中那样,我们不需要假定 R 或 E 是有穷的,但在有穷自动机理论的当前应用中,我们将作出这种限制性假设。

第四个原始概念是在刺激集合上的一个测度概念 μ。假使集合 S 是有穷的,这个测度通常是 S 中的元素的个数。就一般理论而言,我们将假设,S 本身的测度总是有穷的,即 $\mu(S)<\infty$。

第五个原始概念是样本空间 Ω。这个样本空间的每个元素 ω 都代表一次可能的实验,也就是说,试验的一个无穷序列。在当前的理论中,每次试验都可以由一个有序的五元组(C, T, s, r, e)来描述,这里,C 是条件作用函数,T 是在特定试验中为生物体呈现的刺激的子集,s 是 T 的被抽样子集,r 是对试验作出的反应,以及 e 是在那个试验中发生的强化。不可能作出这个

理论的一种全面的解释和理解所要求的全部评论。对于渴望了解更详细描述的那些人，已经给出的两个参考文献将证明是有用的。奈马克（Neimark）和埃斯蒂斯主编的一本论文集（1967），把关于刺激-抽样理论的一组非常综合性的论文收集在一起；也参见埃斯蒂斯的两篇文章（1959a，b）。当前对刺激-反应理论的看法，在许多方面，应被称为刺激-抽样理论，但是，我坚持用更一般的刺激-反应术语来强调，行为心理学的一般思想和语言学理论这两个方面是并列的。此外，刺激-反应理论的特殊的抽样问题，在这里作出的理论应用中不如在实验数据的分析中重要。

因为后来对每次试验中呈现的刺激集合 T 极为重视，所以，应该明确提到它在经典学习理论中的解释。在简单学习的情况下，例如，在经典条件作用下，集合 T 在所有的试验中都相同，我们通常把集合 T 和 S 看成是一样的。在鉴别学习的情况下，集合 T 随试验的变化而变化，我们正在讨论的自动机理论的应用在一般情况下属于鉴别的情形。条件作用函数 C 在反应集合 R 上下定义，C_r 是在这个给定的试验中成为反应 r 的条件或与反应 r 相关的 S 的子集。这些定理解释了条件作用函数如何从一次试验变化到另一次试验。举一个条件作用函数的例子，设 $T = S = \{\sigma_1, \sigma_2, \sigma_3\}$，$R = \{r_1, r_2\}$，和 $C_{r_1} = C(r_1) = \{\sigma_1, \sigma_2\}$ 与 $C(r_2) = \varnothing$。[①]（通常，为了方便起见，用角标符号而不是函数符号表示 C。）在这个例子中，没有刺激成为

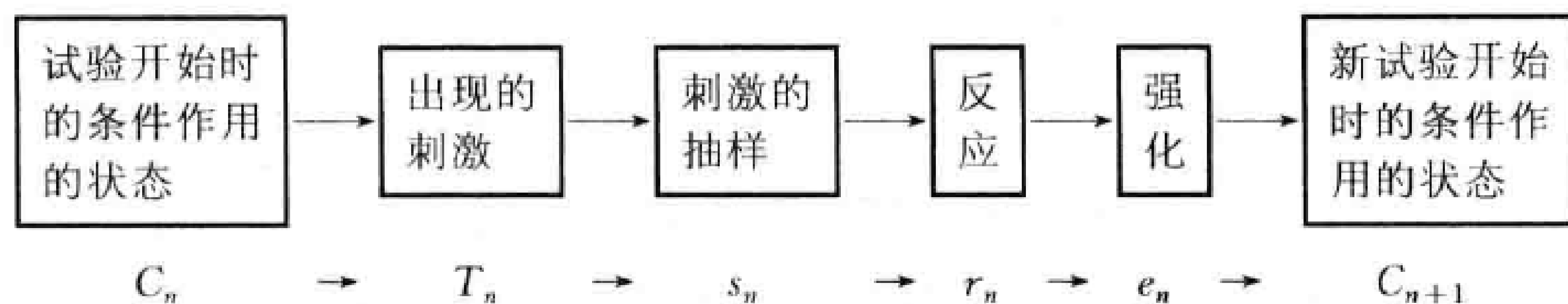
378

① 注意，为了避免歧义，在整个这一节，我用字母"s"只表示被抽样的刺激的集合，因而正如在当前例子中那样，"σ"代表刺激，即 S 的元素。但是，超出这几节仍然有歧义，因为在本章第二节的定义 1 中，用字母"s"表示一个有穷自动机的基本集合 A 的状态。由这种双重用法导致符号可能有歧义之处，在本节只有少数几次，而且，将会被指出来。

r_2 的条件,而且,刺激 σ_3 没有任何条件反射。

从对一个给定试验的这个五元组描述来看,显然,已经作出了关于一次试验中发生的行为的某些假设。特别明显的是,我们假设在一次给定试验中只抽取一个刺激样本,在一次试验中只发生一种反应,也只发生一种强化。这些假设一直是样本空间 Ω 的集合论描述固有的,将不是我们公理的一个明显的组成部分。

用有序五元组表示每次试验的形式的背后是,在任意试验中,从直观上想象事件的时间顺序,可以用下图表示为:



第六个和最后一个原始概念是关于 Ω 的柱集合的适当的 σ-代数的概率测度 P。当刺激集合不是有穷的时,这个 σ-代数的确切描述是相当复杂的,但这种构造是标准的,我们将假设,读者能够从一般的概率理论补充熟悉的细节。值得强调的是,必须根据测度 P 定义所有的概率。

我们也需要某个符号把我们反复带到刺激集合的元素或子集、反应和对样本空间 Ω 的事件的强化之间。首先,r_n 是在试验 n 中的反应 r 的事件,即含有在第 n 次试验中作为一种反应 r 的 Ω 的所有可能的实验实现或元素的集合。同样,$e_{r,n}$ 是在试验 n 中强化反应 r 的事件。事件 $e_{0,n}$ 是在试验 n 中没有强化的事件。依此类推,C_n 是在试验 n 中发生的条件作用函数 C 的事件,T_n 是在试验 n 中发生的呈现集合(presentaiton set) T 的事件,等等。我们将明确指出不遵守这些约定的附加符号。

我们也需要用一个符号表示由按照一个给定试验所发生的

事件所定义的集合。在表达刺激作用和抽样的中心问题与过去事件模式无关时,需要参考这些集合。如果我说,Y_n 是一个 n-柱集,我的意思是说,Y_n 的定义与试验 n 之后发生的任何事件无关。然而,在一次给定的实验中,需要一种更精致的分解来说明假定的序列 $C_n \rightarrow T_n \rightarrow s_n \rightarrow r_n \rightarrow e_n$,因此,在说 Y_n 是一个 C_n-柱集时,意思是说,它的定义与在试验 n 中的 C_n 之后发生的事件无关,即,例如,它的定义可以与 T_{n-1} 或 C_n 有关,但与 T_n 或 s_n 无关。我用简写符号 $Y(C_n)$ 表示这个集合,其他柱集也是如此。如果没有附加限定,符号 Y_n 总是指一个 n-柱集。

379

此外,为了避开太烦琐的符号,我们将使用已经标明的这类事件符号,例如,$e_{r,n}$ 表示在试验 n 中对反应 r 的强化,而且,符号 $\sigma \in C_{r,n}$ 表示在试验 n 中作为反应 r 条件的刺激 σ 的事件。为了简化这些公理的形式陈述,不必重复明显的陈述,假定以概率为条件的任何给定的事件都有正向概率。这样,例如,公理 S2 的默认假设是 $P(T_m) > 0$ 和 $P(T_n) > 0$。

这些公理很自然可分成三类。刺激要成为条件,必须得到抽样,而且,这些刺激必须成为条件,才能产生系统的反应模式。因此,自然有三种公理:抽样公理、条件作用公理和反应公理。我们给出每个公理的口头表述及其形式陈述。从文献中已有的理论表述的观点来看,也许当前公理的最不寻常的特征不是要求刺激集合 S 是有穷的。也应该强调的是,对于任何一种特殊类型的详细应用来说,需要附加专门的假设。这些假设的某种迹象将在自动机理论的特殊应用中体现出来,但这将使我们离题太远,以至于不能在任何细节上并带着对不同实验应用所需要的假设范围的忠诚,来探索这些专门假设。

定义 1. 一个结构 $\mathfrak{S} = (S, R, E, \mu, \Omega, P)$ 是一个刺激-反应模型,当且仅当,满足下列公理:

抽 样 公 理

S1. $P(\mu(s_n) > 0) = 1$。

(在每次试验中,以概率 1 抽样出正测度刺激的一个集合。)

S2. $P(s_m \mid T_m) = P(s_n \mid T_n)$。

(如果在两次不同的试验中产生了同样的呈现集合,那么一个给定样本的概率与试验次数无关。)

S3. 如果 $s \cup s' \subseteq T$ 和 $\mu(s) = \mu(s')$,那么 $P(s_n \mid T_n) = P(s'_n \mid T_n)$。

(成为呈现集合的子集的相等测度的样本,在一次给定的试验中,被抽样的概率相等。)

S4. $P(s_n \mid T_n, Y_n(C_n)) = P(s_n \mid T_n)$。

[给定刺激的呈现集合,在实验 n 中,一个特殊样本的概率与任何前面的事件模式 $Y_n(C_n)$ 无关。]

条件作用公理

C1. 如果 $r, r' \in R$, $r \neq r'$ 和 $C_r \cap C_{r'} \neq \varnothing$,那么 $P(C_n) = 0$。 380

(在具有概率 1 的每一次试验中,每个刺激元素最多成为一种反应的条件。)

C2. 存在 $c>0$,使得对于每一个 σ, C, r, n ,s, e_r 和 Y_n,

$P(\sigma \in C_{r,n+1} \mid \sigma \notin C_{r,n}, \sigma \in s_n, e_{r,n}, Y_n) = c$。

(这个概率是任意被抽样的刺激元素成为强化反应条件的概率 e,假如反应还不是如此有条件的话,而且,这个概率与这个特殊反应、试验次数或任意前面的事件模式 Y_n 无关。)

C3. $P(C_{n+1} \mid C_n, e_{0,n}) = 1$。

(由于概率为 1,如果反应都没有被强化,所有刺激元素的条件作用依然是相同的。)

C4. $P(\sigma \in C_{r,n+1} \mid \sigma \in C_{r,n}, \sigma \notin s_n, Y_n) = 1$。

(由于概率为1,所以,没有被抽样的刺激的条件作用没有改变。)

反 应 公 理

$R1$. 如果 $\bigcup_{r\in R} C_r \cap s \neq \varnothing$,那么

$$P(r_n \mid C_n,\ s_n,\ Y(s_n)) = \frac{\mu(s \cap C_r)}{\mu(s \cap \bigcup C_r)}。$$

[如果至少有一个被抽样的刺激成为某个反应的条件,那么,任何一个反应的概率是成为这种反应条件的被抽样刺激的测度与所有被抽样的成为条件的刺激的测度之比,并且这个概率与任意前面的事件模式 $Y(s_n)$ 无关。]

$R2$. 如果 $\bigcup_{r\in R} C_r \cap s = \varnothing$,那么存在着一个数 ρr,使得

$$P(r_n \mid C_n,\ s_n,\ Y(s_n)) = \rho r\ 。$$

[如果被抽样的刺激都没有成为任何一种反应的条件,那么任何一种反应 r 的概率是一个恒定的推测概率 ρr,概率 ρr 与 n 和任意前面的事件模式 $Y(s_n)$ 无关。]

有穷自动机的表征。分析我们如何能用刺激-反应模型表示有穷自动机的一个有用的开端是,考察最直接的可能进路错在哪里。所出现的困难可以通过两个字母的一个字母表(即两种刺激 σ_1 和 σ_2,还有"启动"刺激 σ_0)和两种状态的一个自动机(即两种反应 r_1 和 r_2)的简单例子来说明。考虑这个例子将有益于后面讨论的几种观点。

凭借公理 S1,在每次试验中必须对所呈现的单一刺激进行抽样,而且,我们假设,对于每个 n,

$$0 < P(\sigma_0 \in s_n),\ P(\sigma_1 \in s_n),\ P(\sigma_2 \in s_n) < 1。$$

进一步假设，这个机器的转换表是： 381

	r_1	r_2
$r_1\sigma_1$	1	0
$r_1\sigma_2$	0	1
$r_2\sigma_1$	0	1
$r_2\sigma_2$	1	0

这要求运用 r_i 和 σ_j 的知识来预测下一个反应应该是什么。模拟这个机器的自然而明显的强化程序表是：

$$P(e_{1,n} \mid \sigma_{1,n}, r_{1,n-1}) = 1$$

$$P(e_{2,n} \mid \sigma_{1,n}, r_{2,n-1}) = 1$$

$$P(e_{2,n} \mid \sigma_{2,n}, r_{1,n-1}) = 1$$

$$P(e_{1,n} \mid \sigma_{2,n}, r_{2,n-1}) = 1$$

这里，$\sigma_{i,n}$是在试验 n 中被抽样的刺激 σ_i 的事件。但对于这个强化程序表来说，这两种刺激中的每一种刺激的条件作用，从一次试验到另一次试验，一直不断变化，正如通过下列序列可能例示的那样。为了简单化和不失一般性，我们可以假设，条件作用参数 e 是 1，而且，我们需要表明没有抽样，如前所述，因为抽样出每个呈现集合中的一个刺激元素的概率为 1。我们可以用 $S=\{\sigma_0, \sigma_1, \sigma_2\}$ 的子集表示条件作用的状态(假定每个刺激都会成为 r_1 或 r_2 的条件)。因此，如果$\{\sigma_1, \sigma_2\}$代表条件作用函数，这意味着，元素 σ_1 和 σ_2 都是 r_1 的条件；$\{\sigma_1\}$意指，只有 σ_1 是 r_1 的条件，等等。然后，考虑从试验 n 到第 $n+2$ 的下列序列：

$(\{\sigma_2\}, \sigma_2 \in s, r_1, e_2) - (\varnothing, \sigma_2 \in s, r_2, e_2) - (\varnothing, \sigma_2 \in s, r_2, e_1)$。

在试验 $n+1$ 中的反应满足这个机器表，但在试验 $n+2$ 中的反

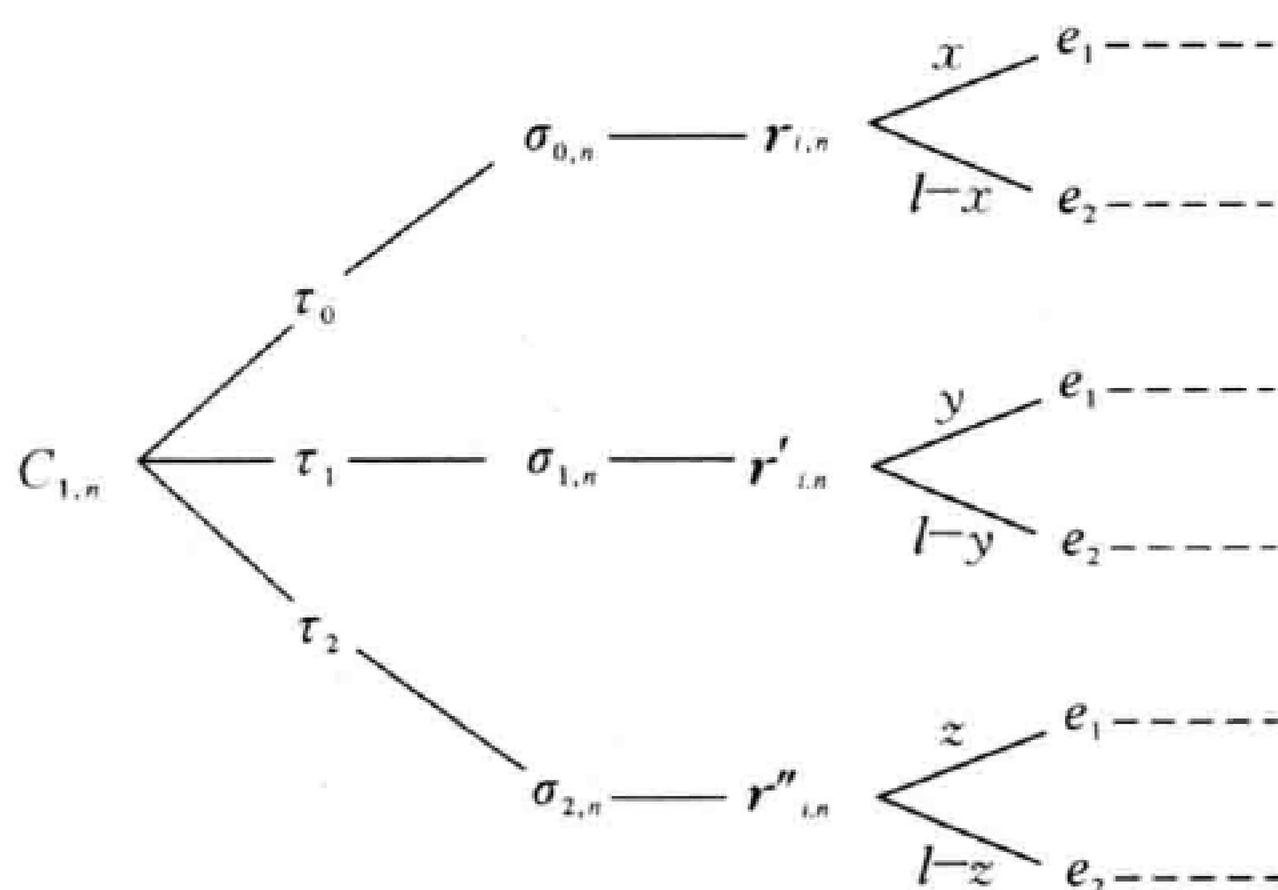
应并非如此,因为 $r_{2,\ n+1}\sigma_{2,\ n+2}$ 后面应该是 $r_{1,\ n+2}$。很容易表明,对于四种可能的条件作用状态中的任何一种状态来说,这种困难是基本的和会出现的。(读者在明确解决这些困难时,应该假设,每一种刺激要么成为 r_1 的条件,要么成为 r_2 的条件,当 n 远远大于 1 和 $e=1$ 时,也是如此。)

所需要的是刺激-反应模型的马尔可夫链的状态的一个相当不同的定义。[关于刺激-反应理论的一般马尔可夫链定理的证明,参见,埃斯蒂斯和苏佩斯的文章(1959b,1974)。]朴素地说,把 S 中刺激的条件作用的可能状态当作马尔可夫链的状态,是很自然的,但在当前的情形中,这在两个方面是错误的。首先,我们必须使反应**模式**和呈现集合成为条件反射,因此,我们把 $R\times S$,即反应集合 R 和刺激集合 S 的笛卡儿积,当作这个模型的刺激集合。生物体必须对试验 n 作出的条件反射是,由在试验 $n-1$ 中给出的前一次反应和在试验 n 中产生的呈现集合组成的模式。

用 $R\times S$ 中元素的条件作用状态定义马尔可夫链的状态,382 仍然是不充分的,因为埃斯蒂斯和苏佩斯(1959b)以及苏佩斯和阿特金森(Atkinson,1960)举了许多例子说明的和明确地给出的理由,有必要把前一次试验中实际产生的反应 r_1 包括在状态定义中。假如 $r_{i,\ n-1}$ 不包括在状态的定义中,所产生的困难可能通过在已考虑过的两种状态的自动机情况下试图画出的树型图显示出来。假设只有模式 $r_1\sigma_1$ 是有条件的,其他四种模式 σ_0, $r_1\sigma_2$, $r_2\sigma_1$①和 $r_2\sigma_2$ 是没有条件的。让我们用 C_1 表示这个条件作用状态,并设在每次试验中,Γ_j 是 σ_j 的必然概率,其中,$0\leqslant j\leqslant 2$,每个 $\tau_j>0$。那么这个树型图看起来是这样的:

---

① 原文有误,在与作者沟通后改为现在的字母。——译者



这个树型图并没有完成，因为如果不知道在试验 $n-1$ 中实际发生的反应，我们就无法完成这些分支（例如，详细说明这些反应），由于同样的原因，我们不能确定概率 x，y 和 z 。此外，我们无法通过把在试验 $n-1$ 中的可能反应包括在这些分支中，来挽回这种局面，因为我们要确定它们的概率，就需要考虑试验 $n-2$，这种回归直到我们回到试验 1 时才能结束。

因此，我们把在试验 $n-1$ 中的反应包括在状态定义中。另一方面，在确定型有穷自动机的情况下，没有必要在这些状态中允许 $R\times S$ 中这些模式的所有可能的条件作用。我们将只允许两种概率——这个模式是没有条件的，或者，对于适当反应来说，它是有条件的，因为发生错误反应的条件作用的概率为零。因此，由于内部状态或反应的概率是 p 和在 Σ 中有 m 个字母，所以，有 $(m+1)p$ 种模式，其中每一种模式都处于两种状态之一，有条件的或无条件的，以及有 p 种可能的前反应，因此，在马尔可夫链中的状态数是 $p2^{(m+1)p}$。实际上，方便的是，进一步把 σ_0 看作单一模式来简化这个数，不管它与什么样的前反应成为一对。于是，状态数是 $p2^{mp+1}$。因此，对于最简单的 2 种状态、2 个字母表的自动机来说，状态数是 64。我们可以用有序的

$mp+2$ 元组表示这些状态：

$$(r_j, i_0, i_0, 1, \cdots, i_0, m, \cdots, i_{p-1, m}),$$

这里，$i_{k,l}$是 0 或 1，取决于模式 $r_k\sigma_l$ 是无条件的，或是有条件的，其中，$0 \leqslant k \leqslant p-1$ 和 $1 \leqslant l \leqslant m$；$r_j$ 是前一次试验中的反应，i_0 是 σ_0 的条件作用状态。我们希望证明的是，在完全没有条件的状态$(r_j, 0, 0, \cdots, 0)$的集合中，从概率 i 开始，这个系统将最终总是处于这样一种状态中，即这种状态是全部有条件的状态$(r_j, 1, 1, \cdots, 1)$的集合的一项。证明这一点是证明基本表征定理的主要部分。

383

在转向这个定理之前，我们需要明确地定义逐渐成为一个自动机的一个刺激-反应模型的概念。正如已经建议的那样，这是这个定义的一个重要特征。与自动机的字母表 Σ 相对应的基本的刺激集合 S，不是这个刺激-反应模型的基本的刺激集合，更确切地说，这个基本集合是笛卡儿积 $R \times S$，这里 R 是反应的集合。[①] 此外，在每个试验中只允许出现和抽样 S 的一个元素，以这样一种方式，已经表达了该定义；然而，为了概念和符号的简化，这里所用的限制是无关紧要的。如果没有这种限制，基本集合就不是 $R \times S$，而是 $R \times p(S)$，其中，$p(S)$是 S 的幂集，即 S 的所有子集的集合，于是，这个字母表 Σ 中的每个字母都是 S 的一个子集，而不是 S 的单一元素。重要的是使 $R \times S$ 而不是 S 作为基本的刺激集合。

例如，对(r_i, σ_j)必须被抽样和作为一种模式的条件，而且，阐述这些公理要求在一次给定试验中所抽样的和所成为条件的是呈现集合 T 的一个子集。在这种联系中，为了简化符号，我

---

① 在下文中牢记这个区别是重要的。注意，为了有助于强调这一点，后面两段进一步讨论用 S_R 代替 R。

通常写成 $T_n = (r_{i,\,n-1},\ \sigma_{j,\,n})$，而不是写成

$$T_n = \{(r_i,\ \sigma_j)\},$$

但意义是清楚的。T_n 是由在试验 $n-1$ 中的反应 r_i 构成的单一模式(或元素)与在试验 n 中的刺激元素 σ_j 组成的呈现集合，而且，从公理 S1 来看，我们知道，抽样出这个模式，是因为它是所呈现的惟一模式。

运用 $R\times S$ 在形式上是方便的，但根本没有必要。分析的经典 S-R 传统提供了形式上等价的但在心理学上更现实的一条进路。每个反应 r 都显示了一种刺激 σ_r，或更一般地说，显示了一个刺激集合。再次假设，为了形式上的简明性，只有一个刺激元素 σ_r，而不是一个刺激集合，根据完全偶然的呈现程序表，我们可以用刺激集合 S_R 代替 R，

$$P(\sigma_{r,\,n} \mid r_{n-1}) = 1,$$

而且，在这个模型中，我们现在考虑笛卡儿积$(S_R \times S)$，而不是 $R\times S$。在这个框架内，关于每次试验中的呈现集合的重要一点是，一个分量是纯粹的可控的受试者，另一个分量是纯粹的可控的实验者——如果我们用熟悉的实验区分的话。在训练动物像自动机一样执行任务时，约定明确地运用 S_R，而不是 R，是重要的，因为额外引入 σ_r 直接而有意义地减轻了动物的记忆负担。对于儿童语言学习的模型来说，S_R 的重要性还不太明显。

本着本章最后一节的精神，从心理学过渡到神经科学能够给出这个理论的另一种解释。这种过渡意味着用反应的大脑表征取代反应本身。另一方面，这个在生物学上更现实的过渡将不改变本节基本论证的形式结构，但一种经验上现实的说明将会推出第 6 节给出的那种类型的更多细节。 384

定义 2. 设 $\mathfrak{S} = (R\times S,\ R,\ E,\ \mu,\ \Omega,\ P)$ 是一个刺激-反应

模型，这里

$$R = \{r_0, \cdots, r_{p-1}\},$$

$$S = \{\sigma_0, \cdots, \sigma_m\},$$

$$E = \{e_0, \cdots, e_{p-1}\},$$

以及对于 $S' \subseteq S$，$\mu(S')$ 是 S' 的基数。于是，$\mathfrak{S}$ 渐进地成为自动机 $\mathfrak{A}(\mathfrak{S}) = (R, S - \{\sigma_0\}, M, r_0, F)$，当且仅当，

(i) 当 $n \to \infty$ 时，概率为 1，所以，对于某个 i 和 j，呈现集合 T_n 是 $(r_{i,n-1}, \sigma_{j,n})$。

(ii) $M(r_i, \sigma_j) = r_k$，当且仅当，对于 $0 \leqslant i \leqslant p-1$ 和 $1 \leqslant j \leqslant m$，

$$\lim_{n\to\infty} P(r_{k,n} \mid T_n = (r_{j,n-1}, \sigma_{j,n})) = 1。$$

(iii) 对于 $0 \leqslant i \leqslant p-1$，

$$\lim_{n\to\infty} P(r_{0,n} \mid T_n = (r_{i,n-1}, \sigma_{0,n})) = 1。$$

(iv) $F \subseteq R$。

正如已经评论的那样，关于这个定义的次要的但可起到澄清作用的一点是，刺激 σ_0 不在自动机 $\mathfrak{A}(\mathfrak{S})$ 的字母表内，因为在初始状态 r_0，需要把一种刺激强加给这个自动机，从这里提出的理论观点看，这要求生物体将作出反应 r_0 所需要的一种刺激。这种刺激就是 σ_0。这个定义也要求，从渐进的趋势来看，刺激-反应模型 $\mathfrak{S}$ 只是自动机 $\mathfrak{A}(\mathfrak{S})$。应该清楚的是，一个更弱和更一般的定义是可能的。自动机 $\mathfrak{A}(\mathfrak{S})$ 只能逐渐地被嵌入到 $\mathfrak{S}$ 当中，并且只是 $\mathfrak{S}$ 的活动的一部分。达到这种概括的最简单方式是，使这个自动机的字母表成为 $S - \{\sigma_0\}$ 的惟一的一个真子集，与之相应的是，构成自动机内部状态的那些反应；它们只需要反应的全集 R 的一个真子集。这里不再关注这种概括，尽

管此类问题对于给出语言的语义方面的充分说明是必要的。

定理 1. (有穷自动机的表征定理)**给出任何一个连通的有穷自动机,存在着一个逐渐地变得同构于它的一个刺激-反应模型。此外,这个刺激-反应模型拥有的所有反应可能最初都是无条件的。**

证明: 设 $\mathfrak{A} = (A, \Sigma, M, s_0, F)$ 是任何一个连通的有穷自动机。正如已经指出的那样,对于 $0 \leqslant i \leqslant p-1$,我们用反应 r_i 的集合 R 表示内部状态集合 A,这里,p 是状态数。我们用刺激 $\sigma_1, \cdots, \sigma_m$ 的集合表示字母表 Σ,并且,由于已经明确的原因,我们把 σ_0 增加到这个刺激集合中,获得

$$S = \{\sigma_0, \sigma_1, \cdots, \sigma_m\}。$$

为了后面参考,设 f 是在 A 与 R 之间和 Σ 与 $S-\{\sigma_0\}$ 之间建立一一对应的 $A \cup \Sigma$ 上定义的函数。(为了回避一些不重要的技术要点,我假设 A 和 Σ 不相交。)

我们把下列集合看成是强化集合: 385

$$E = \{e_0, e_1, \cdots, e_{p-1}\}$$

并且,对于 $S' \subseteq S$,测度 $\mu(S')$ 是 S' 的基数,因此,像定义 2 中那样,我们正在考虑一个刺激-反应模型 $\mathfrak{S} = (R \times S, R, E, \mu, \Omega, P)$。为了表明 $\mathfrak{S}$ 渐进地成为一个自动机,我们对 $\mathfrak{S}$ 强加五个附加限制。这些限制如下。

第一,在强化 e_0 的情况下,这个程序表是这样的:

$$P(e_{0,n} \mid \sigma_{0,n}) = 1 \tag{1}$$

即如果 $\sigma_{0,n}$ 是试验 n 中的呈现集合的一部分,那么强化反应 r_0 的概率为 1——注意,强化事件 $e_{0,n}$ 的概率与反应 $r_{0,n}$ 的实际发生无关。

第二,其余的强化程序表用自动机 $\mathfrak{A}$ 的转换表 M 来定义。明确地说,对于 j, $k \neq 0$ 和所有的 i 与 n, $P(e_{k,n} | \sigma_{j,n} r_{i,n-1}) = 1$,当且仅当,

$$M(f^{-1}(r_i), f^{-1}(\sigma_j)) = f^{-1}(r_k)。 \tag{2}$$

第三,对这个证明来说,重要的是超越(1)和(2)的附加假设:刺激 σ_0, $\cdots$, σ_m 中每一个刺激都有一个正向概率,即在每次试验中,必然发生的概率(我们能构造出一个含有更弱假设的模型,但是,弱化这个必要条件是没有意义的)。明确地说,我们因而假设,对于任何一个柱集 $Y(C_n)$,使得 $P(Y(C_n)) > 0$,对于 $0 \leqslant i \leqslant m$ 和所有的试验 n,

$$P(\sigma_{i,n}) = P(\sigma_{i,n} \mid Y(C_n)) \geqslant \tau_i > 0, \tag{3}$$

第四,我们假设,当没有抽样出成为条件的刺激时,发生反应 r_i 的概率 ρ_i 也严格为正,即,对于每个反应 r_i

$$\rho_i > 0 \tag{4}$$

这强化了公理 $R2$。

第五,对于每个整数 k, $0 \leqslant k \leqslant mp+1$,我们把集合 Q_k 定义为恰好有 k 个条件模式的状态集,$Q_{k,n}$是在这样一种状态中的事件:这个状态是在试验 n 中的 Q_k 的一个项。我们假设,在试验 1 开始时,所有的模式都是无条件的,即

$$P(Q_{0,1}) = 1 \tag{5}$$

很容易证明,已知集合 R、S、E 和基数测度 μ,存在着满足限制(1)—(5)的许多不同的刺激-反应模型,但要证明这个定理,没有必要选择模型类的某个特征项,因为下面的论证表明,这个类的所有项都逐渐地变得同构于 $\mathfrak{A}$。

我希望证明的主要问题是

$$\lim_{n\to\infty} P(Q_{mp+1,\ n}) = 1。\tag{6}$$

我们首先注意到，如果 $j < k$，那么从 Q_k 转换到 Q_j 的概率为零，即

$$P(Q_{j,\ n} \mid Q_{k,\ n-1}) = 0 \tag{7}$$

此外，

386

$$P(Q_{j,\ n} \mid Q_{k,\ n-1}) = 0 \tag{8}$$

即使 $j>k$ 成立，除非 $j=k+1$。换言之，在一次试验中，至多一种模式能够成为有条件的。

为了表明，从渐进趋势来看，(6)成立，只要证明：存在一个 $\epsilon>0$，使得在每次试验 n 中，当 $0\leqslant k\leqslant mp<n$ 时，如果 $P(Q_{k,\ n})>0$，那么

$$P(Q_{k+1,\ n+1} \mid Q_{k,\ n}) \geqslant \epsilon \tag{9}$$

为了确立(9)，我们需要证明，一种刺激模式的概率至少是 ϵ，这个刺激模式在试验 n 开始时是无条件的，后来在这个试验中变成有条件的。所给出的这种论证将是对任意无条件的模式都成立的一致论证。设 $r^*\sigma^*$ 是在试验 n 中的这样一种模式。

现在，众所周知，就一个连通自动机而言，对于每一种内部状态 s，都存在着一个纸带 x，使得①

$$M(s_0,\ x) = s \tag{10}$$

而且，x 的长度小于等于内部状态数。根据刺激-反应理论，x 是刺激元素的长度不大于 p 的一个有穷序列。因此，我们根据 $\sigma_{i_p}=\sigma^*$，可以得到 $x=\sigma_{i_1}$，…，σ_{i_p}。我们根据(3)知道，

① 在等式(10)中，我们在本节有少数情况之一：这里，"s"和"s_0"指被模拟的自动机的状态，不是指被抽样的刺激的集合。

$$\min_{0\leqslant i\leqslant m}\tau_i = \tau > 0 \tag{11}$$

所要求的反应序列 r_{i_1}，…，r_{i_p-1}，要么从前面的条件作用来看是会产生的，要么如果具有猜想概率 r_i 的任何一种反应，相对于适当的模式都是无条件的，也是会产生的，借助于(4)

$$\min_{0\leqslant i\leqslant p-1}\rho_i = \rho > 0 \tag{12}$$

为了表明，在试验 n 中，模式 $r^*\sigma^*$ 有一个正向条件概率 ϵ，对于"运行"的纸带 x 来说，我们只需取 n 足够大，比如说，$n>p+1$，并且考虑联合概率，

$$P^* = P(\sigma_n^*,\ r_{n-1}^*\sigma_{i_p-1,\ n-1},\ \cdots,\ r_{0,\ n-i_p-1}\sigma_{0,\ n-i_p-1}) \tag{13}$$

这个基本公理和假设(1)—(5)确定了一个与 n 无关的 P^* 的下界。首先，我们注意到，根据(3)和(11)，对于刺激元素 σ_0，σ_1，…，σ^* 中的每个刺激元素，

$$P(\sigma_n^* \mid \cdots)\geqslant\tau,\ \cdots,\ P(\sigma_{0,\ n-i_p-1})\geqslant\tau$$

同样，从(4)和(12)还有反应公理来看，我们知道，对于反应 r_0，α_{i_1}，…，r^* 中的每个反应，

$$P(r_{n-1}^* \mid \cdots)\geqslant\rho,\ \cdots,\ P(r_{0,\ n-i_p-1} \mid \sigma_{0,\ n-i_p-1})\geqslant\rho$$

因此，我们知道

$$P^* \geqslant \rho^p\tau^{p+1}$$

387 而且，已知发生的事件 $\sigma_n^* r_{n-1}^*$，条件作用的概率是 c，据此，我们可以获得

$$\epsilon = cp^p\tau^{p+1} > 0$$

这确立了(9)，证明结束。

已知刚才证明的这个定理，就有几个有意义的推论，这些推

论的证明几乎是直接的。第一个推论把正则语言的表征定理与要产生的有穷自动机的表征定理结合起来:

定理 2. 任何一种正则语言都是由某个渐进的刺激-反应模型生成的。

一旦引入概率的考虑,我们就能以几种不同的方式超越刺激-反应产生的语言的局限推延至正则语言的范围。本节后面就这样做。

我猜想,在这里给出的意义上,愿意接受把有穷自动机和正则语言简化为刺激-反应模型的许多心理学家或哲学家,很少对下列主张感到满意:我们同样能简化意图、计划和目的概念的明确意义。然而,我在这方面没有有价值的新的分析,由米勒(Miller)和乔姆斯基(1963)已经作出的分析,能够做到这一点。故事是这样的。1960 年,米勒、加兰特尔(Galanter)和普里布拉姆(Pribram)出版了书名为《计划和行为结构》的一本引发争议的书。在这本书中,他们严厉地批评刺激-反应理论能说明的人类和高级动物的有意义的行为是如此之少。他们尤其反对把条件反射作为建立一门适当的科学心理学的一个恰当概念。在我的印象中,许多认知心理学家都感到,在这本书中对 $S-R$ 理论的批评是毁灭性的。

正如我在前面指出的那样,我同意,条件反射实验确实太简单,不能成为分析更复杂行为的一种适当的科学基础。我们举例来说,这就像从初等集合代数推导微分方程理论的企图一样没有希望。然而,更一般的集合论确实在一种严格的数学意义上包含了微分方程理论。

可以证明,就米勒和乔姆斯基系统地表达的后面这个理论而言,在刺激-反应理论和计划理论(theory of plans)之间也有同样的关系。米伦森(Millenson,1967)非正式地提出了这一

点。计划理论是根据“tote”单元理论(“tote”是由测试-运算-测试-退出这个周期的第一个字母缩写而成的)来表达。于是,把一个计划定义为一个 tote 层次结构,这只是一种形式的定向图,而且,每一个有穷定向图都可用一个有穷自动机来表示。因此我们有下面的结论:

定理 3. *在米勒和乔姆斯基的意义上的任何一个 tote 层次结构都同构于某个渐进的刺激-反应模型。*

对于批评的回应。[①] 阿尔比布(Arbib,1969)对有穷自动机的刺激-反应理论的评论,提出了继续把认知心理学家和刺激-反应的心理学家区分开来的问题。尽管我愿意承认,复杂的学习和行为的完全适当的刺激-反应理论还有待发展,但是,他对下列观点的任何论证都说服不了我,这种观点是:心理学家迄今为止已经定义了对刺激-反应理论的一个切实可行的替代选择。阿尔比布的评论和论证能够从下列五个方面作出分析。我依次加以研究,并在几种情况下,也详述了我自己的观点。

388

证明的问题。阿尔比布宣称,他已经更简单地推导出我的主要定理:已知任何一个连通有穷自动机,存在着一个刺激-反应模型渐进地同构于它。他通过假设“对于每个刺激-反应表来说,都有一个逐渐地同构于它的刺激-反应模型[p. 507]”简化了这个证明;但是,正如很久以前伯兰特·罗素所言,用这样的假定步骤代替证明,是选择了避难趋易。换言之,阿尔比布并没有表明,基本结果是如何从关于刺激-反应联结的简单的一般假设中得出的——我的主要目标。从问题的技术层面来看,他面临的问题是表明,他的很有效的假定(对应于我的框架中的定理)与关于条件作用和刺激抽样的简单假设是一致的,这种条件作

① 这一节的其余部分来自苏佩斯(1969c,1989)。

用与刺激抽样是任何一个标准的刺激-反应理论的一个内在组成部分。

刺激痕迹和中介反应。 一个生物体如何能记住它的过去的反应,阿尔比布似乎对这个问题持有相当简单的观点。他提到了在通过一个适当的刺激使前面的反应外在化的动物实验中用到的自然方法。在表现出像简单自动机一样的鸽子的条件作用中,这条进路恰好得到了运用,但它不是一般的进路。无疑,这不适合于我自己的研究最初提出的学习语言的问题,乃至儿童学习算术的问题。

在经典的刺激-反应理论中,自然要谈到刺激痕迹,我本来就提到过这种可能性。依照后来的刺激-反应理论,尤其是与奥斯古德(Osgood)、马尔茨曼(Maltzman)、伯莱因(Berlyne)、肯德勒斯夫妇(the Kendlers)等人相关的那些理论,谈论生物体内部的中介反应(mediating response)是自然的。这两个概念中任何一个概念都提供了替代或扩展阿尔比布所用的很狭义的外在化进路的一个自然框架。他反对在有机体内有对前面反应的某种表征,以免"如果输入通道必须既传送外界刺激的表征,也传送有关 TOTE 层次结构的当前执行状态的完整信息,输入通道将会完全超载[p. 509]"。然而,这个评论,像阿尔比布的大多数其他评论一样,也不会受到详细分析的支持。大家都不明白,这是一个问题。以更一般的方式来说,人们最好也对内部处理表示关注,并且仅仅说,已经发现生物体到底如何能够工作,是一件过分神秘的事情。

状态数。 我认为,在许多方面,阿尔比布提出的最认真和最重要的问题是质问:一个生物体究竟是否可能有充分的时间学习,它要成为一个适当大小的自动机所需要的大量明显的刺激-反应联结。这特别是由批评语言学习的刺激-反应理论的心理

语言学家不断地提出的一个核心问题。阿尔比布对含有八个分量每个分量有四种状态的 TOTE 层次结构作出的否定陈述,是熟悉的,他的论证也假设了一种熟悉的神学论调,说那简直是无法办到的。

因为这些年来我对这些问题已经有许多论证,所以,我对否定的信条与否定的证明作出区分。许多心理语言学家和认知心理学家断言否定的信条:刺激-反应思想无法说明复杂的学习,因为表现出复杂行为的生物体绝对没有足够的时间学习适当的条件作用联结。站在包括阿尔比布在内的几乎所有断言这个信条的人的立场上,这个否定的信条似乎是一种信念,而不是一个证明。无疑,他没有证明:一个有 4^8 种状态的计算机是他喜欢定义的任何一个认知过程所必需的。从他熟悉的关于这些问题的自动机理论的正式文献来看,我怀疑,在冒险给出一个详细的论证似乎是什么时,他多少比许多心理语言学家更慎重。关于与一个否定的信条相反但实际上从头到尾一片混乱的一个假定的否定证明的例子,参见由贝弗(Bever)、福多尔和加勒特(Garrett,1968)所建议的证明:刺激-反应理论或联想主义的一个形式的限制,是能够被证实的。

在数学及其相关的学科中一般承认,否定的论证必须以比肯定的论证更形式和更明显的方式来阐述。在数学和哲学中,为评价否定论证的有效性,确立明确的和认真定义的系统标准,也有很长的历史。似乎对我来说,阿尔比布并没有在这种经典传统中工作。他作出过关于状态数的一些非正式的陈述,但并没有认真提供关于下列认知过程所要求的自动机大小的否定论证:这些认知过程已经得到了很好的理解,足以用相对精确的术语加以描述。我随后将再回到这一点。

我认为,不同的任务所要求的状态数问题确实是一个基本

问题,我当然不准备为完成极其大量的任务所需要的状态数确定很小的上界。另一方面,我对心理语言学家或自动机理论家如此经常作出的大量陈述表示怀疑。我只提及两个例子。在苏佩斯的一篇文章的结尾(1969b),我详细地分析了两个十位数相加所需要的自动机类型,我表明,这种自动机如何能够以现实的方式与学生的实际行为相关。① 所需要的状态数恰好是两个。不经过详细分析的非正式想法可能会建议说:状态数将需要是大的。事实并非如此。

我不打算详述的第二个例子是,努力写出两三岁孩子的口语的概率文法。尽管我不希望主张,这些努力是完全成功的,但从小孩子说出的句子长度来看,显然,这种文法的产生式所要求的状态数是相对小的,也许,只有处理词汇才需要大的状态数。 390 我应该补充说,我还没有遇到任何一位语言学家或自动机理论家希望主张,能够通过与学习的直接经验过程相当不同的任何其他方式获得词汇。我推测,包括语言理解或产生式在内的任何一项认知任务所要求的状态数,原来证明比我们当初想到的要小得多。我愿意承认,这只是一种推测,但我有足够的信心坚持认为,由于问题是状态数必须是有条件的,所以,愿意主张刺激-反应理论是不充分的那些人,应该提供否定的证明,而不只是否定的信条。我自己努力研究的核心焦点是,把我自己的肯定信条转化为肯定的证明。(关于状态数的问题,或许值得注意的是,只用七种状态就能构造一个通用图灵机。)

用一个简单的例子来说明在谈论状态数时的一个重要观点。如果我们把一枚硬币投掷一百次,可能的结果数是 2^{100}。我们每个人都能做这个实验,然而,我们即使能让全世界的人都

① 本节前面也举过这个例子。

投掷硬币,也不可能逼近可能的结果数。重要的是注意到,就等级或数量级来说,一个过程中的实际状态数通常比可能状态数小得难以令人置信,条件作用联结的可能数和实际数之间的关系也是如此。

真学习与隐喻学习。阿尔比布在几个地方认为,按照刺激-反应理论,学习简直不可能发生,因为所要求的状态数太大,而且,条件作用所需要的时间太有限。我已经表达了我的中心论证,反对他关于状态数的主张。我将愿意认同,如果各个必须是有条件的状态数很大,那么完成任务的时间将不够。阿尔比布所考虑的替代选择的歧义性和缺陷,是令人感兴趣的。按照我的主要定理的第一推论,他明显表示出对 TOTE 层次结构的偏爱,没有把这些层次结构考虑成是刺激-反应模型的特殊情况。他没有说出乃至勾画出,在 TOTE 层次结构的框架内,如何能给出生物体获得认知技能的一个理论。他确实偶尔提及有关大脑神经元和生物化学变化的某些参考文献,但是,在这个生理学框架内没有提出认真的想法。他没有告知读者,阿尔比布所建议的认真的替代选择是什么,一种例外的可能是,他最终把计算机作为一个隐喻之脑进行评论。大概隐喻之脑从事隐喻学习。如果这就是他所建议的学习理论,那么他不能意指大家认真地采纳了他在这个方向上的评论。对生物体如何学习的说明,必须以十分明确的方式,用完全非隐喻的术语来陈述,因为它要受到实验的检验。无疑,阿尔比布本人对如何解决机器学习的理论或人类学习的理论面临的任何一种主要困难,并没有提出肯定的或明确的建议。

确实,令人惊讶的是,依照他关于心理学的学习理论的适当性的见解,他没有更多地谈到计算机的学习理论(即机器学习)的优点。然而,我以为,理由实际上是相当明显的。不需要对人

391

工智能的文献进行很认真或系统的研究，就能得出这样的结论：尽管心理学的学习理论无疑还很不完善，但是，应该如何为计算机编写学习程序的理论，也是如此。①

层次结构问题。我已经陈述过，阿尔比布提出的状态数问题确实是一个重要的问题。我认为，他提出的另一个重要问题是层次结构问题。对于几乎所有的认知心理学家来说，条件作用理论必须把每个简单的条件作用联结看成是分离的和平等的，这似乎是一个信念问题。似乎所建议的层次结构概念与刺激-反应思想的精神正相反。我的文章中的主要定理是，米勒和乔姆斯基意义上的每个 TOTE 层次结构，都渐进地同构于某个刺激-反应模型，这个定理的推论直接表明，存在着一个明显的理论反例。这个抽象性结果可能不会令某些读者感到满意，但它的意义是明确的。在形式上可以通过一个 TOTE 层次结构、一个有穷定向图、一个有穷树来表征或一个有穷自动机来表征的任何一个直观的层次结构，恰好也可以通过一个刺激-反应模型来表征；事实上会更好，因为如果刺激-反应联结还没有在基因中得到编码，那么刺激-反应模型也包括对如何了解这种层次结构的一种说明。

对表征定理的批评。基拉斯(Kieras,1976)对定理 1(即有穷自动机的表征定理)从技术上提出了详尽而有趣的批评。在基拉斯关于上述定理太强的主张中，很容易鉴别出基拉斯混淆的直观原因。② 因为我认为给定自动机的*内部状态*等同于表示刺激-反应模型的*反应*，所以，基拉斯推断，我无意识地把我自己

---

① 在机器学习语言甚至是学习自然语言的一个简单片段的理论中，需要什么样的细节，将在本章第 5 节作出很好的举例说明。这不是建议说，那里所用的特殊进路是惟一的进路。反正当前很难看出，如何才能避免许多明确理解的细节。

② 安德森的批评本质上基于同样的混淆。

的分析局限于在它们的内部状态和反应之间具有一一对应的自动机。基于他的这种混淆，他断言，表征定理不像它看起来那样是正确的。

我现在的目的是，直截了当地揭示这种争论，以便明确地表明，基拉斯是错误的，最初陈述的上述表征定理是正确的。从数学的观点来看，基拉斯的错误在于误解了表征定理。表征定理的同构是一种形式上的同构。在上述定理1的情形中，同构是自动机的内部状态和表示刺激-反应表征模型的反应之间的同构。在我的1969年的文章中用到的雷宾-斯科特定义的自动机，没有明确的反应机制，但是，第2节关于带输出的有穷自动机的定义2，以及哈里森(Harrison,1965, p. 294)给出的带输出时序机的一般定义表明，在雷宾-斯科特定义中附加反应机制是不重要的，基拉斯参考了哈里森的定义，并且，哈里森承认，他主要采用了雷宾和斯科特的术语(1959)。一个带输出的自动机或时序机，对于哈里森来说，只是在具有下列附加条件的雷宾和斯科特意义上的一个自动机：最终状态的集合 F 是“给出一个输出”的那些状态。正如哈里森等人所评论的那样，限于输出1和0并不是对时序机的一般性的限制。

392

把这种输出方法附加到我当初文章里给出的形式定义中是没有价值的。我只选出不用来表示内部状态的两种反应 r_0 和 r_1，但其中一种反应，比如说，r_0 表示0，另一种反应表示1。只要机器所在的内部状态不是最终状态，但务必作出一种反应，它就输出 r_0。当机器处于最终状态时，它输出 r_1。为了考虑这些输出反应来修改定义1并不困难。我再一次指出，这两种输出反应不打算对应于所表示的自动机的内部状态。刺激-反应模型的其他反应表示这些内部状态。我也强调，增加输出反应的这种修正，将不会纠正定理1中的错误，而只会提供一个附加的

密切相关的结果。

确定型强化的问题。正如几位批评者所注意到的那样，也正如我在几种场合[例如，在苏佩斯文章中(1975)]所评论的那样，在多种学习环境中，与证明定理1的要求相反，反应的严格确定的强化是不可行的。例如，如果我问一个小孩子7+5等于多少，那么，当他给出一个不正确的答案时，一种确定型强化将给出正确答案12。一种不确定型强化的例子只是告诉他，回答不正确，请他重试。但在运用确定型强化时，显然，在某种意义上，为了纠正每一个不正确的反应，这些反应必须是可观察的。定理1运用了一个确定型强化假设。似乎对我来说，正是确定型强化的假设，而不是我关于可观察的反应或视为输出的内部状态作出的任何一种非形式的评论，是这种结果的核心限制。

从为复杂学习提供一种说明的立场来看，尤其是在与简单的实验室情境相对的自然环境中，我很清楚，在定理1发表之前，最重要的扩展是用非确定型强化获得类似的结果。我与我以前的学生威廉·罗特迈尔(William Rottmayer)联合解决了这个问题，详细结果体现在他1970年的学位论文中。对这种结果的一种非形式的和相当简要的陈述出现在我们发表的关于自动机的一篇概述性文章中(Suppes and Rottmayer,1974)。

因为具有非确定型强化的刺激-反应理论的正式陈述相当复杂，我只给出类似于苏佩斯和罗特迈尔的文章(1974)中简要的非形式陈述。在这样做之前，我先明确阐述，在消化各个定理的内容时，有直观参考价值的一个公认的问题类。把刺激显示的一个潜在的无穷类呈现给生物体，例如，线图。整个类的一个子集通过一个有穷自动机来描述。学习者的问题是，已知在每次试验中，仅有的信息是，他对给定的刺激显示的分类是否正确，据此，学习由有穷自动机描述的概念。我这里已经提到线

393

图,因为我不想完全集中于语言,但也有可能根据识别某一正则语言的文法字符串来思考学习。因为我不喜欢完全根据文法来思考语言学习,在当前语境中,我更喜欢一个几何学的事例。让我们假设,每个线图都是由有限的几条线段组成的。一个典型的例子是,由三条线段组成的线图,构成了一个三角形,但是,其中一条线段并且只有一条线段能延伸到三角形之外。在每次试验中,都要求学习者说出,所示的特殊线图是否是这样一个图形,在他说出答案后,只告知他,他的答案是否正确。

这个理论所假设的是,学习者为了对看到的图形作出分类,在说出"是"或"否"的答案之前,已经作出的一系列的潜在反应或内部反应(如果你喜欢的话)。非形式地假设了潜在反应或内部反应是不可观察的或不能被直接强化的。这个理论不要求把这一点作为一个假设,但在这个理论的任何一个实验应用中隐含了这一层意思。强化不是出现在每一个内部反应发生(被看成是一个子试验)之后,而是只出现在由一系列子试验组成的一个试验之后。

换言之,用标准的实验术语来说,一个子试验对应于我们通常当作的一次试验,但是,没有强化或条件作用发生,我们无法观察到实际作出的反应。条件作用只有在一系列子试验之后才出现,整个系列的子试验被称为一次试验。用自动机的术语来说,一个子试验对应于一个自动机进行了一次转换,也就是说,从一个内部状态到另一个内部状态,而且,一次试验对应于处理整个纸带或输出字符串。

条件作用出现在有一个正确反应的试验中,去条件作用出现在有一个错误反应的试验中。这样,学习发生在所有的试验中,与反应是否正确无关。以这些公理为基础,能够证明下列定理,这个定理代表了对我 1969 年文章中的基本定理的很大改

进。这种改进归功于强化方法的弱化。

定理 4. *如果 D 是感知显示的任意集合，G 是能够被一个有穷自动机所识别的 D 的一个子集，那么，存在一个也能学习识别 G 的刺激-反应模型，这个模型表现出渐进地与自动机的表现相一致。*

需要注意的重要的一点是，由于非确定性强化，这个定理正如人们所预期的那样是较弱的。在确定型强化的情况下，刺激-反应模型渐进地变得同构于给定的有穷自动机。在当前的情况下，结果只能是一种行为等价，或者，用通常的自动机理论的语言来说，结果是一种弱等价。

显然，在我刚阐述的理论中运用的非确定型强化是对令人感兴趣的强化的最弱的看法，一种可能的例外是只给出部分强化，也就是说，在某些实验中的强化。在实际学习时，例如，在学习数学或学习语言时，有许多情形能给出更确定的和更有益的强化。不难表明，在一般情况下，对于有特定能力的生物体而言，强化越确定，学习就越快。在强化概念的长期复杂的发展史上，人们一直不太强调，强化是传递信息，信息的特殊结构隐含在任何一个特殊的强化框架内。对不同强化框架进行的一种详尽但不全面的分析，可在贾米森(Jamison)、拉蒙(Lhamon)和苏佩斯(1970)的文章中找到。(这篇文章作出了许多详细的理论计算，以至于很少有人能读懂；当运用含有略微复杂的信息结构的强化框架时，问题会迅速变得何等复杂，这篇文章对这个问题确实提供了一种很好的理解。) 394

在当前的讨论中重要的是，考虑一种最弱的非确定型强化结构，并强调这种观点：正是强化的非确定型特征使我们走出反应的经典的可观察性范围，允许引入通常不可观察的潜在反应或内部反应的一个指令表。这种强化不会产生潜在反应，但

当我们有了确定型强化时，这个理论在发生了潜在反应或内部反应的情境中是不可应用的。

同样重要的是注意，不管这些潜在反应是被称为这样的反应，还是被称为内部状态，这其实是一个术语问题，不是一个实质性的理论问题。很容易重新表述上面给出的公理，而且，除了在一个试验结束时发生的反应之外，也很容易用内部状态替代反应。这种术语上的变化不会以任何方式影响这些公理。

值得提到的是，我们可能希望命名为内部状态的潜在反应，甚至在我们不知道它们的确切形式时，通常也被看成是会发生的。一个很好的例子是，描述大多数沉默的成人读者的默读发音反应的情况。这种默读反应的自我意识是不同寻常的，我得赶紧补充说，通常不可能从默读反应“读出”正被读的词。

另一种误解：限于有穷自动机。文献中的另一种误解是，刺激-反应理论只能研究具有有穷自动机能力的机器。本小节的目的是，通过构造等价于图灵机的寄存器机表明，情况并非如此。这里的发展实质性地扩展和修改了苏佩斯的文章(1977b)。为了赋予这些结果形式上的明晰性，我们将针对任何一个部分递归函数提出一个学习理论。我能够以相当直接的方式定义这样的函数，但是，这里我不会这么做。我将依赖的事实是，部分递归函数是可计算的函数。我们于是运用了文献中的基本定理：任意函数是部分递归的，当且仅当，通过一个寄存器机或一个等价的图灵机，它是可计算的，我们稍后将会详述这个定理的技术构架。这里所用的一个寄存器机的概念是在第3.5节中引入的。用寄存器机而不用图灵机的理由是，寄存器机的形式结构较简单。例如，一个函数是部分递归函数和是通过一个寄存器机可计算的函数之间的等价性证明，比有关图灵机的相应证明更加简单。首先，让我从3.5节回想起，对于一个

395

有穷词汇表来说,一个经典的寄存器机是多么简单。我们大家都有一个潜在的无穷列表或寄存器序列,但是,任何一个给定的程序只能用一个有穷数。确切地说,对于每个寄存器来说,需要三条简单的指令。第一条指令是,把这个有穷词汇表的任何一个元素都放到寄存器 n 的内容顶部;第二条指令是,如果这个寄存器是非空的,就删除寄存器 n 的内容的最下面的字母;因为任何一个计算都是在有限的步骤内完成的,所以,任意寄存器的内容在长度上一定总是有穷的。第三条指令是,如果寄存器 n 的内容最下面或开始字母是 a_i,就跳到程序的另一行的一个跳转指令;换言之,这是一个条件跳转指令。因此,如果我们把寄存器的内容看成是从左到右的读取字符串,我们也能把这些指令描述为在右边放置新符号,在左边删除旧符号,而且,如果需要,在寄存器中运用一个条件跳转指令。

对于这样一种无限寄存器机来说,给出一个程序的形式定义是简单明了的,但我当前暂时不这样做。一个程序显然只是由刚才描述的那类几行指令组成的。就寄存器数和每个寄存器的大小而言,一个无限寄存器机的潜在的无穷记忆,是一种自然的数学的理想化。也有可能的是,定义带有刚才陈述的那类指令的单寄存器机,并表明一个单寄存器也是适当的。

这里给出的修改刺激-反应理论的一个重要观点是,编码刺激显示所用的内部语言全部是经过处理的。换言之,在当前阐述的寄存器机理论中,我将不考虑外部刺激集合和编码语言之间的关系,而只研究这种显示的已编码的表征。这种抽象度,对于当前的讨论来说,似乎是适当的,但当然,对于一个很完善的理论来说,还是不合适的。然而,根据适当的模式化,这是一个合适的劳动分工。我正在假设,由于在中枢神经系统之外很好地出现了第一层次的编码,所以,感觉系统把这种已编码的信息

传达到中枢神经系统。因此，在某种意义上，刺激概念就这样变成了非功能性的，但只是因为已经假设了编码。很显然，关于刺激的实际感知特征的重要假设，不是经典的 $S-R$ 理论的一个组成部分。第二，用内在地构造的一个程序概念取代以刺激为条件的反应的直接语言。一个自然的问题是，为什么不试着更多地从神经系统层面来提供这种构造。已知我们对信息传到中枢神经系统的实际方式所知甚少，然后，用来编码和编程，这似乎是不成熟的。无疑，情况似乎是，存在着内部编程。我不是在建议说，一个寄存器机的简单的抽象理论抓住了内部编程的细节——这只是表示内部编程的一种方式——一个详细的附加的理论问题是，修改抽象的表征使它变得更现实。

396 另一方面，如果没有给出多少有点详细的神经系统的分析的话，寄存器机的程序能被在计算上等价的刺激-反应联结取而代之，但是，如果不进一步作出详细的说明，这种假定的 $S-R$ 条件作用联结就不会比寄存器机的程序更具体，即更接近于经验实现。因此，似乎对我来说，把这些程序看成是由我们目前还不能具体说明的神经"硬件"来实现的会更好些。本节的其余部分所呈现的内容是很形式化的，但无疑在许多方面有待改进，要么更接近地模仿不同生物体的学习，要么使机器学习更有效。此外，给出所呈现的刺激的某种编码特征，就有好的理由认为，对于解决本质上具有相同学习速率的一个特殊的问题类来说，任何一个软件程序都存在着一个相应的神经网络，反之亦然。

为了使问题更明显和更形式，但不试图完全形式化，我想到了在 3.5 节特别是在任意有穷字母表上的递归函数那一小节的下列定义。首先，<n>是寄存器 n 在执行一条指令前的内容；<n'>是寄存器 n 在执行一条指令后的内容。第二，一个寄存

器机有(1) 一个可数的寄存器序列,编号为1,2,3,……其中,每个寄存器都能存储来自有穷字母表 V 的符号的任何一个有穷序列,(2) 三条基本指令是:

(a) $P_N^{(i)}(n)$: 把 a_i 放在 $<n>$ 的最右边。

(b) $D_N(n)$: 如果 $<n> \neq \varnothing$,删除 $<n>$ 最左边的字母。

(c) $J_N^{(i)}(n)[q]$: 如果 $<n>$ 从 a_i 开始,跳转到行 q。①

如果跳转到一个不存在的行,那么机器停止。指令中所示的下角标参数 N 是指这样的特征寄存器的集合:即它包含有感觉数据并且不能用作正常计算的寄存器。(在下列给出的定义中,使这一点更加明显。)

一个寄存器机的程序的某一行,要么是由一个大于等于1的自然数 m(行数)和指令(a)或(b)之一组成的一个有序数偶,要么是由一个大于等于1的自然数 m、指令(c)之一和一个大于等于1的自然数 q 组成的一个三元组。这个定义的直观解释显而易见,将不再给出。

(一个寄存器机的)一个**程序**是一个 k 行的有穷序列,使得(1) 第 i 行的第一个数是 i,(2) 每行的第三项都是数 q,其中,$1 \leqslant q \leqslant k+1$。参数 k 当然是这个程序的行数。我也把程序看成是**例行程序**。一个寄存器机如何**遵循**一个程序或例行程序在直观上是明显的,不再从形式上加以定义。像程序一样定义**子程序**,下列情况除外:(1) 子程序可能有几个退出,(2) 三元组的第三项可能在范围 q_1, …, q_k 之内——在一个给定的程序中为这些变量赋值。

在字母表 V 上定义的一个部分递归函数的形式定义,在

① 这里对3.5节的符号稍微做了一点改动,用"q",而不是"$E1$",表示所跳转到的行。这是为了方便后面给出证明。

397 3.5节结束时给出。我简单重复其要点。这样一个函数是任何一个直观上可计算的函数。已知 V,即有穷词汇表,于是,照例在这些问题中,V^* 是 V 的元素的有穷序列的集合;在当前语境中,我将把 V^* 的元素称为特征编码。设 f 是从 $V^* \times \cdots \times V^*$ (n 次)到 V^* 的 n 个自变量的一个函数。基本定义是,f 是通过一个寄存器机可计算的,当且仅当,对于每一个寄存器 x_i, y 和 N,其中,$y \neq x_i, i=1, \cdots, n$ 和 $x_1, \cdots, x_n, y \leqslant N$,存在一个例行程序 $R_N(y=f(x_1, \cdots, x_n))$,使得如果 $<x_1>, \cdots, <x_n>$ 是寄存器 $x_1, \cdots, x_n$ 的初始内容,那么

1. 如果 $f(<x_1>, \cdots, <x_n>)$ 是未定义的,机器将不会停止;

2. 如果定义了 $f(<x_1>, \cdots, <x_n>)$,机器将停止于 $<y>$,寄存器 y 的最终内容等于 $f(<x_1>, \cdots, <x_n>)$,而且,除了 y 之外,所有的寄存器1、2……,N 的最终内容都与开始时一样。

寄存器学习模型的公理。我现在转向寄存器学习模型的公理,广而言之,它们与在苏佩斯和罗特迈尔文章(1974)中为具有非确定型强化的刺激-反应模型提供的那些公理相并列。我只使这个模型公理化,不使作为学习试验的一个形式框架的整个概率空间公理化。可能通过随机变量和使概率空间成为潜在的来扩展后者,这点是直截了当的,但单调乏味。

这些公理基于下列结构概念:

(i) 寄存器的集合 R,

(ii) 模型的词汇表 V,

(iii) 特征寄存器的子集 F,

(iv) 计算寄存器的子集 C,

(v) 反应寄存器的子集 Rp,

(vi) 工作记忆 WM,

(vii) 长期记忆 LTM,

(viii) 反应 r_0 和 r_1,

(ix) 实参数 p 和 c。

也许有用的是,简要地和非形式地讨论每个原始概念。F 中的特征寄存器只编码所呈现的刺激的特征。运用有穷词汇表 V 进行这种编码和计算。C 中的计算寄存器是在需要计算时可用的工作寄存器。工作记忆 WM 存储所构造的程序。为了简单起见,这里我将假定,只有一种这样的记忆,但显然,为了达到一般的目的,这太有限。长期记忆 LTM 是通过重复正确试验发现的程序所在的地方。

在两种记忆和这些寄存器之间的一种区分,是基本的。这两种记忆存储程序,因此,V 中的特征词汇 v_1, …, v_n 是三条指令的附加符号:P 表示放在或加在右边,D 表示删除左边,J 表示跳转指令。词汇表 V 也必须包括指所用的寄存器和程序行的符号。为了达到这个目的,我增加了单位数字 1(这样,2=11,3=111,等等),即最起码的计数符号。

398 为了简单起见,这里也假设,反应寄存器的集合 Rp 是一个单元素集合。在前面描述的通用寄存器机中,这个寄存器对应于保存所计算的部分递归函数的数值的寄存器。一个无关紧要的简化假设是,学习将局限于概念学习,这原则上根本不限于可计算函数的集合。在当前的情况下,已知程序已经完成,如果清空寄存器,反应是 r_0,这意味着,所显示的刺激——它的特征在 F 中编码——是一个要学习的概念事例;如果寄存器是非空的,反应是 r_1,这意味着,所呈现的刺激不是一个要学习的概念事例。此外,如果程序在完成之前停止于任何一步,反应是 r_0 的猜测概率为 p,是 r_1 的概率为 $1-p$。

两个实参数 p 和 c 以相当不同的方式进入这些公理当中。正如刚才所表明的那样，p 是反应的猜测概率，c 是停止程序构造的常概率。这些参数和在这些公理中潜在地引入的其他参数，当然是依赖于语境的，很自然，会随着任务的变化而变化。

正如这里所阐述的那样，运行一个程序的每一行，都被认为是选入程序构造并置入工作记忆(WM)当中。一个程序，只有在它结束后，成功地正确识别了一个要学习的概念事例，才能被转换到长期记忆(LTM)当中。这种完全清除所构造的错误程序的机制太苛刻，但却是对某些动物学习成立的一种简化假设，例如，海兔通过感官刺激具有的全或无(all-or-none)的淘汰习性(Kandel, 1985)。

前面提到的寄存器机的这三条指令——右边添加，左边删除，或条件跳转——在一个方面作了修改。跳转到不存在的一行，因此而停止程序，而不是跳转到 $m+1$，这里 m 是行数，跳转到 0，即 0 是无行的一个可能数字。这种改变的理由应该是明显的。当程序是由学习模型在概率意义上逐行地构造时，就无法预先知道程序将会有多长。因此，为了停止程序，方便的是，预先有一个固定的"地方"来跳转。

定义 3. 一个结构 $\mathfrak{R}=(R, V, F, C, Rp, WM, LTM, r_0, r_1, p, c)$ 是一个概念形成的寄存器学习模型，当且仅当，满足下列公理：

寄存器结构公理

$R1$. 寄存器的子集 F、C 和 Rp 是非空的且两两不相交。

$R2$. 子集 F 与 R_p，以及集合 V，是有穷的和非空的。

$R3$. R 中的每个寄存器都能保存 V_1^* 的任何一个词，即 $V_1=V-\{1, P, D, J\}$ 的元素的任何一个有穷字符串。

刺激编码公理

$D1$. 在每个试验开始时，在 F 的寄存器 f 中，把所呈现的刺激编码为具有特征$\langle f\rangle$。

程序构造公理

$P1$. 如果在试验开始时，LTM 是非空的，则程序构造不会发生。

$P2$. 已知 LTM 是空的：　399

(i) 在 WM 中以概率 $c(0<c<1)$的程序构造在每行之后都会终止，与试验次数和任意前面的事件子序列无关；

(ii) 已知程序增加了一行，任何一个自变量的任何一条指令的抽样概率都是正的，与试验次数和事件前面的任意子序列无关；在行数 n 的情况下，发生跳转的概率呈几何分布。

程序执行公理

$E1$. 如果 LTM 是非空的，则把内容复制到 WM，然后，执行程序。

$E2$. 如果 LTM 是空的，那么，按照构造公理 $P1$ 和 $P2$，在概率意义上逐行地构造一个程序，并且，每构造一行，执行一个程序。

$E3$. 在执行一个跳转指令时，一步后程序以一个固定的正向概率停止，这个概率与试验次数和事件前面的任意子序列无关。

反应公理

$Rp1$. 如果当程序是完备的时，寄存器 Rp 是空的，则反应是 r_0。

$Rp2$. 如果当程序是完备的时，寄存器 Rp 是非空的，则反应是 r_1。

$Rp3$. 如果根据公理 $E3$ 程序停止，则作出反应 r_0 的猜想

概率为 p,作出反应 r_1 的概率为 $1-p$;概率 p 与试验次数和事件前面的任意子序列无关。

强 化 公 理

*Er*1. 如果在一个试验结束时发生了正强化,则假如 *LTM* 是空的,*WM* 中的程序就被复制到 *LTM* 中。

*Er*2. 如果在一个试验结束时发生了负强化,则 *WM* 中的程序就被消除,而且,假如 *LTM* 是非空的,*LTM* 中的程序也是如此。

需要对一些公理作出评论,这些评论在前面的非形式讨论中是没有的。概率程序构造公理 *P*2 类似于这样一个刺激抽样公理:它确保了所有相关刺激的条件作用的可达性。公理 *P*2 显然是以这样一种方式阐述的:例如,限制抽样概率不渐近地趋于零,抽样行数呈几何分布的情形除外。为了防止其余的程序生成无限循环,程序执行公理 *E*3 中要求的停止概率是必要的。最后,这些公理中用到的非正式的强化概念具有明显的意义,并容易被形式化。这里的正强化只意味着,根据反应 r_0 或 r_1 对一种刺激的概念分类是正确的,而负强化意味着这是不正确的。显然,更能提供信息的强化方法能够并广泛地用于学习中,而且,无疑提高了学习速度。在最后关于层次结构学习的评论中,会更多地讨论这种观点。

400 基于上面陈述的这些公理,我们可以证明,一个通常对应于刺激-反应模型的定理 1 的渐进学习定理。

定理 5. 设 f 是有穷字母表 V 上的 n 个自变量的任意部分函数,并且,在 V 中只有两个值。于是,f 是一个部分递归函数,当且仅当,f 是通过关于概念形成的一个寄存器学习模型 $\mathfrak{R}$ 以概率 1 逐渐地可学会的。

证明: 设 p 是计算 f 的 $\mathfrak{R}$ 的一个程序。我们知道,一个有

穷字母表上的一个函数 f 是部分递归的，当且仅当，它是通过一个寄存器机可计算的，根据这个事实，一定存在着这样一个程序。而且，已知 f 的定义，我们有产生 $\mathcal{P}$ 的一种构造方法。我们的目的是表明，在由这些公理描述的学习环境中，在每次试验中都存在着一个构造 $\mathcal{P}$ 的正向概率。

设 $C \subseteq V^* \times \cdots \times V^*$（$n$ 次）是 f 的已编码的刺激实例的集合——可计算的概念 C——不失一般性，在这种语境中，我把这个概念看成是它的实例的集合，再设 $\neg C$ 是 C 的补集。我们把刺激分布看成是一种呈现，这里，$(<f_1>, \cdots, <f_n>)$ 是一种刺激的特征编码表征，

$$P((<f_1>, \cdots, <f_n>) \in C)$$

$$= P((<f_1>, \cdots, <f_n>) \in \neg C) = \frac{1}{2}。$$

此外，我们以下列几何方式设计从 C 和 $\neg C$ 中抽样的实验。设 f_i 是刺激 σ 的特征 i 的在 V^* 中的编码。再设 $|f_i|$ 是在 f_i 中的符号数。于是，$\sum |f_i|$ 是编码 σ 所用的符号总数。我们用一个几何分布表示符号总数，用一个均匀分布表示在符号总数相同的那些符号中的选择。（在一个完全形式化的理论中，关于所呈现的刺激的概率选择的这些假设将是下列公理的一个组成部分，即我这里只限于寄存器学习模型的公理，在任何细节上都不包括关于刺激呈现或强化程序的公理。）

现在假设，最初 LTM 是非空的。如果存储在 LTM 中的程序能正确地计算 f，我们就完成了任务。如果对于某些刺激 σ 来说，这个程序不能正确地计算 f，那么，根据刚才陈述的假设，在每次试验中，将以一种固定的正向概率呈现 σ，因而，根据公理 $Er2$，以概率 1 逐渐地清空 LTM。

于是，在每次试验中，构造 p 的概率都是正的。详细的计算是这样的：首先，设 p 有 m 行。根据公理 $P2$(i)，构造一个只有 m 行的程序的概率等于 $c(1-c)^{m-1}$。如果行 i 不是一个跳转指令，那么，根据公理 $P2$(ii)，具有期望形式的行 i 的概率大于某个 $\epsilon_1 > 0$。如果行 i 是一个条件跳转指令，在这里，跳转到行 n_i，那么，同样根据公理 $P2$(ii)，对于某个 $\epsilon_2 > 0$，行 i 正好是程序 $\mathcal{P}$ 的行 i 的概率等于 $\epsilon_2^2(1-\epsilon_2)^{n_i-1}$。

因此，这些概率的有穷积在每次试验中都是正的，与试验次数无关。显然，设 $i_1, \cdots, i_{m_1}$ 是非跳转指令的行，设 $j_1, \cdots, j_{m_2}$ 是跳转指令的行，其中，$m=m_1+m_2$，那么

401

$$\mathcal{P}\text{的概率} > \epsilon_1^{m_1} \Pi_{i=j_1}^{i=j_{m_2}} \epsilon_2^2(1-\epsilon_2)^{n_i-1}, c(1-c)^{m-1} > 0 \tag{I}$$

根据这个不等式，我们立刻推出，将以概率 1 逐渐地学习 $\mathcal{P}$，这就完成了证明，下列评论除外：所构造的程序描述了一个部分递归函数，证明这一点是简单明了的。

层次结构的作用和更加确定的强化。对于寄存器-模型概念学习的理论来说，正如定义 3 所阐述的那样，我们不能改进不等式(Ⅰ)。如果把(Ⅰ)看作一个等式，显然，对于任意长度的程序 $\mathcal{P}$ 来说，学习将是非常缓慢的，比我们在大多数人类的学习甚至大多数动物的学习中所看到的学习还要慢得多。

在当前的理论框架内，在一个合理的时段发生学习的惟一实际希望是，把学习分解为所掌握的相对小的任务的一个层次结构。可能有人认为，通过使强化比上述公理 $Er1$ 和 $Er2$ 中所假设的内容提供更多的信息或更加确定，就可以避免这个结论。这种观点有一定的正确性和重要性，它能得到详细计算的重要例子的支持。另一方面，也有一个解释的问题。对于在证明定理 1 时用到的完全确定的强化来说，我能够把有穷自动机的每

个内部状态的条件作用都看成是一项任务——这里，任务是通过强化的内容来定义的，根据这种观点，最细致化的层次结构是由完全确定的强化创造的。

为了表明能使理论回到现实当中并应用于数据，有用的是，以把理论应用于小的熟悉的任务而告终。当然，在当前的语境中，我将不打算认真处理实际参数的估计。所选择的任务是，当呈现出三角形、四边形和五角形时，让五岁的孩子通过识别三角形来学习三角形的概念。

关于这些孩子所用的寄存器模型，我作出下列假设。(在这些问题上，它有必要的简化类型。)

(i) 语言 V_1 有单一元素 α，它是用于计数的。

(ii) 有两个特征寄存器，#1 表示线段数，#2 表示大小，其中，α=小，$\alpha\alpha$=中等，以及 $\alpha\alpha\alpha$=大。

(iii) 条件跳转，要么跳转到前一行，要么跳转到 0(表示一个不存在的行和停止)。

(iv) 为了简化表述，在特征寄存器上进行直接计算，而不是首先把特征寄存器的内容复制到工作寄存器。(根据三条原始指令描述从一个寄存器到另一个寄存器的复制，是简单明了的。)

(v) Rp 是单反应寄存器。

(vi) 设 a 是选择删除指令的概率，b 是选择跳转指令的概率，$1-a-b$ 为选择放置或增加指令的概率。

(vii) 设 p 是选择特征寄存器 1 的概率，$1-p$ 是选择参照程序的一行的特征寄存器 2 的概率。

一个简单的正确程序是： 402

1. D(1)从寄存器 1 删除 α。

2. D(1)从寄存器 1 删除 α。

3. D(1)从寄存器 1 删除 α。

4. Copy(1, Rp),把寄存器 1 的内容复制到反应寄存器 R_p 中。

这里所用的缩写形式的所有程序,必须以把特征寄存器或工作寄存器的内容复制到反应寄存器中而告终。于是作出一种反应。因此,第 1—3 行的概率是: $p^3a^3c(1-c)^2$,其中,c 是公理 $P2$(i)中引入的行数分布的参数。

重要的是承认,许多不同的程序将产生正确的反应,因此,一种正确反应的概率远大于 $p^3a^3c(1-c)^2$。如果任务是识别四边形,而不是三角形,那么,即使对所考虑的简单实验情境而言,作出全面的分析也是非常复杂的。尽管如此,在合理假设的条件下,接近最短长度的正确程序的概率应该控制着一种正确反应的理论计算。

这里在公理意义上定义的学习方案,就其范围而言,比得上部分递归函数的定义,或者,计算这些函数的寄存器机的定义,即这些定义适用于分别考虑的每一个函数。但对概念层次结构的延伸学习而言,为了达到实际的学习速率,必须利用以前学过的概念来丰富这个结构。这里举一个非常简单的例子来说明这一点。考虑由 n 种不相交的情形构成的析取概念。只需要一个寄存器,字母表 V_1 是集合{α, β},没有跳转指令,不过只有四个指令表示在左边删除字母或在右边增加字母。每种情形的程序最多有 10 行。于是,假定对指令和行数(1 到 10)的抽样有一个均匀分布,能直接计算出最多有 10 行的每个程序的概率。更重要的是,我们能够容易计算可能的程序数:长度为 1 的有 4 个,长度为 2 的有 16 个,一般情况下,长度为 n 的有 4^n 个,其中,$1 \leqslant n \leqslant 10$,在当前的例子中,总数为$(4^{11}-4)/3$,这约等于 4^{10}。现在如果在程序的第二阶段把只用原始指令与个别情形

的 n 个子程序和准许的长度不超过 $2n$ 的程序加在一起，那么就有 $[(n+4)^{2n+1}-(n+4)]/(n+3)$ 个可能的程序，这约等于 $(n+4)^{2n}$。另一方面，如果设计一个程序在一步内有 $10n$ 行，则可能的程序数约等于 4^{10n}。例如，考虑 $n=3$ 的情况。那么 4^{30} 比 $7^6+3(4^{10})$ 大许多个数量级。这个例子的细节并不重要，我也不打算整理它们足以在两条进路的每一条进路中确定正确的可能程序数。通常，在层次结构进路和非层次结构进路中，这个数占总数的百分比很小。从层次结构进路获得的利益是很明显的。

更一般地说，比如，动态地改变使用前面定义的概念的概率，即它的识别程序的某些聪明的方式，对实际的机器学习至关键要，而且，关于这些方法的可靠假设，对于一个延伸的概念层次结构的任何一种复杂的人类学习或动物学习，似乎都是根本的。同样重要的是，引入了比定义 2 假设的简单情形更丰富的信息反馈形式，但对替代方案的数学研究似乎仍然处于初级阶段——只有极端的情形得到了相对较好的理解。许多人类的学习依赖于口授和纠正，不过，从根本的立场看，这种反馈过程的一个近似适当的理论迄今还没有。各种极其简化的假设，例如，协议分析中运用的那些假设，在当前，似乎是无法消除的。这是对还有多少事情要做的一个衡量尺度。 403

8.4 通过刺激-抽样模型学习的线性模型表征①

为了表明根据一个理论对另一个理论的表征，即使从概念

① 本节专业性很强，在不失连续性的前提下可以跳过。它通过模型的同构确实提供了把一个理论还原为另一个理论的真实事例。

上看是相对简单的,也可能是何等错综复杂和有技术性,我这里证明关于两个密切相关的学习理论的一个定理。这个定理及其证明来自许多年前一篇未发表的技术报告(Estes and Suppes 1959b)。诺曼(Norman,1972)曾发表过关于这个定理及其证明的一个较简单但有限的看法。诺曼的限制很苛刻。限定只依赖于发生在产生强化的相同试验中的事件强化程序表。

与给出把经典热力学还原为统计力学的真正令人满意的证明的复杂性相比,这里考虑的例子是简单的,在数学上是初步的。后面的这个还原当然具有重要的科学价值,但可能还没有充分意识到,为这种还原提供一个适当理论发展有多么复杂,以及获得一个完全令人满意的数学证明必须强加什么样的限制。①

当前的简单例子,就其本身而言,已经是复杂的和技术性的,它提供了在科学的任何一个领域内,根据适当的表征定理,把一个科学理论有意义地和令人满意地还原为另一个理论的某些困难的感觉。重要的是注意到,把一个科学理论还原为另一个科学理论的问题,在概念上完全不同于证明一个测量理论或几何理论的一个表征定理的问题。在后一种情况下,理论模型的结构和特殊的表征模型之间的关系是简单的和明显的。的确,通常的测量理论或几何理论是由一个想到的非常确定的数值解释建构的。这些公理在很大程度上受这种预期的数据解释的引导。把一个科学理论还原为另一个理论通常是在一个相当不同的概念框架内进行的。

本节我们描述了在现有文献(Bush and Mosteller, 1955)中大量讨论的各种线性模型。大多数线性模型应用于简单的学

① 与吕埃勒(Ruelle,1969)关于统计力学中的相应情形的评论相比较。

习实验，即刺激的呈现集合不随试验的变化而变化的那些实验。 404
在埃斯蒂斯和苏佩斯的文章(1959a)中，基于一个学习参数 θ 相当广泛地分析了这些模型的特定类。因而，在 Ω 中的每个 ω 都表示一个试验的无穷序列。ω 的每个试验都用一个有序对 (r_i, e_k) 来表示，其中，$r_i \in R$，即可能反应的有穷集合，$e_k \in E$，即强化的有穷集合。设 ω_n 是一位特定的受试者通过试验 n 的反应和强化序列。在埃斯蒂斯和苏佩斯文章(1959a)中为线性模型所提供的这些公理，假定了已知有穷序列 ω_n 的反应概率的线性关系。

定义 1. 一个结构 $(R, E, \Omega, P, \theta)$ 是一个简单学习的（单一参数的）线性模型，当且仅当，Ω 是序列 ω 的抽样空间，P 是 $\mathfrak{B}(\Omega)$ 上的一个概率测度，即柱集的 σ-代数，学习参数 θ 位于半开区间 $(0<\theta\leqslant 1)$，并且，对于每个 n，Ω 中的每个 ω，R 中的每个 r_i，以及 E 中的 e_k，满足下列三个公理：

公理 1. 如果 $P(e_{i,n}r_{i',n}\omega_{n-1}) > 0$，那么

$$P(r_{i,n+1} \mid e_{i,n}r_{i',n}\omega_{n-1}) = (1-\theta)P(r_{i,n}) \mid \omega_{n-1}) + \theta。$$

公理 2. 如果 $P(e_{k,n}r_{i',n}\omega_{n-1}) > 0$，$k \neq i$，和 $k \neq 0$，那么

$$P(r_{i,n+1} \mid e_{k,n}r_{i',n}\omega_{n-1}) = (1-\theta)P(r_{i,n} \mid \omega_{n-1})。$$

公理 3. 如果 $P(e_{0,n}r_{i',n}\omega_{n-1}) > 0$，那么

$$P(r_{i,n+1} \mid e_{0,n}r_{i',n}\omega_{n-1}) = P(r_{i,n} \mid \omega_{n-1})。$$

注意 e_0 代表没有强化的事件。

本节由四部分组成。第一部分像上一节陈述的那样对刺激-抽样理论的一般公理略微作了修改。第二部分确立了一些预备定理。第三部分阐述了包括反应 r_i 和强化 e_j 的实际序列在内的各个定理。最后，第四部分引入了某些特殊极限假设，并

证明了一般的表征定理。

一般公理的修改。现在回到刺激-反应模型,令人向往的是,用一个符号表示以在试验 n 中一个特定反应为条件的元素的个数,即抽样元素的个数等。一般情况下,如果 A 是任意集合,我们用 $N(A)$表示 A 的基数。沿用在学习理论中的用法,我们设 $N=N(S)$,这里 S 是刺激的集合。此外,$N(s_n)$是 Ω 中所有序列 ω 的集合,在第 n 次试验中正好有 $N(s)$个被抽样的刺激元素。我们以同样的方式运用符号 $N(T)$, $N(T_n)$, $N(C_i)$, $N(C_{i,n})$, $N(s_i)$, $N(s_{i,n})$。需要明确地强调 $N(s)$和 $N(s_n)$之间的区别,等等。$N(s)$是表示在 s(S 的一个子集)中有穷个元素的一个整数排列。相反,$N(s_n)$是抽样-空间点的一个无穷集。像在上一节中一样,我们仍然继续使用

T, T', T'', …表示刺激的呈现集合,

C, C', C'', …表示条件作用函数,

s, s', s'', …表示被抽样的刺激的集合。

运用明确的符号 $N(s)$,我们也通过下列方式加强了前面的公理:即总是使集合 s 成为有穷的,而且,总是把表示 S 的任意子集 s 的 $\mu(s)$解释为 s 中的元素个数,因此写成 $N(s)$,而不是 $\mu(s)$,并保留 $\mu(i)$表示某个量 i 的概率的平均值。

405

我们也需要一种**实验者的划分**(experimenter's partition)的符号,埃斯蒂斯和苏佩斯的文章(1959a)中已经引入了这个符号。我们为这样一种划分提供一个语言定义。对试验 n(含有元素 η, η', η'', …)的一种实验者的划分 $H(n)$是一个抽样空间 Ω 的一种划分,使得 $H(n)$的每个元素 η 只根据事件 $T_{n'}$、$r_{i,n''}$、$e_{j,n'''}$下定义,其中,$1\leqslant n', n'', n'''\leqslant n$。应该清楚的是,一种实验者的划分 $H(n)$的元素只是 n 维柱集的特类。

强化公理

$E1$. 一种强化的概率只取决于前面的可观察量，即前面呈现的刺激集合、前面的反应和前面的结果。

存在一种实验者的划分 $H(n)$，使得如果 W_{n-1} 是任意 $n-1$ 柱集，$Y \subseteq W_{n-1} \cap C_n \cap s_n$，$\eta \in H(n)$，并且，$P(Y \cap \eta) > 0$，那么

$$P(e_{j,n} \mid Y \cap \eta) = P(e_{j,n} \mid \eta)。$$

呈现公理

$P1$. 一个刺激呈现集合的概率只依赖于前面的可观察量。

存在一种实验者的划分 $H(n-1)$，使得如果 W_{n-1} 是任意 $n-1$ 柱集，$Y \subseteq W_{n-1} \cap C_n$，$\eta \in H(n-1)$，并且 $P(Y \cap \eta) > 0$，那么

$$P(T_n \mid Y \cap \eta) = P(T_n \mid \eta)。$$

预备定理。我们确立了许多定理和推论，它们对于主要定理来说是初步的。因为我们限于简单的学习，所以，本节的所有定理都基于下列假设：

（Ⅰ）刺激集合 S 是每次试验中的呈现集合。

根据（Ⅰ），我们可以在下文中忽略呈现集合 T 的所有考虑。为了获得独立于 n 的结果，我们也假设，给定空样本 s^O，反应 i 的概率是 ρ_i，即对于每个 i 和 n，

（Ⅱ）$\rho_{i,n} = \rho_i$。

第三，作为一个符号问题，我明确地把被抽样的刺激数的平均值定义为

（Ⅲ）$\bar{s} = \sum_{N(s)=1}^{N} N(s) P(N(s))$

前五个定理的推论比定理本身更清楚地表明了这些定理如

何为推出线性模型提供基石。前两个定理因试验 n 的反应而异。接下来的两个定理改变了试验 n 的强化事件。第五个定理处理试验 n 的无强化情形。

406 在第一个定理中，我们需要的事实是，下列超几何分布的二阶原始矩：

$$\sum_{a=0}^{A} a^2 \frac{\binom{A}{a}\binom{N-A}{s-a}}{\binom{N}{s}} = \frac{1}{N}\left\{As + \frac{A(A-1)s(s-1)}{N-1}\right\}$$

定理 1. 如果 $P(e_{i,n}r_{i,n}N(C_{i,n})) > 0$，那么

$$P(r_{i,n+1} \mid e_{i,n}r_{i,n}N(C_{i,n})) = \rho_i P(s^O) + (1-P(s^O))。$$

$$\left[\frac{N(C_i)^2}{N^2} + \frac{\bar{s}N(C_i)}{N^2} - \frac{N(C_i)}{N^2} - \frac{N(C_i)(N(C_i)-1)(\bar{s}-1)}{N^2(N-1)}\right]$$

$$\Big/\left[(1-P(s^O))\frac{N(C_i)}{N} + \rho_i P(s^O)\right]$$

证明： 通过通常的级数展开方法，我们有

$$P(r_{i,n+1} \mid e_{i,n}r_{i,n}N(C_{i,n}))$$

$$= \sum_{s=s_i}^{N}\sum_{s_i=0}^{C_i} P(r_{i,n+1} \mid e_{i,n}r_{i,n}s_{i,n}s_nC_{i,n})P(e_{i,n} \mid r_{i,n}s_{i,n}s_nC_{i,n})$$

$$\frac{P(r_{i,n} \mid s_{i,n}s_n)P(s_i \mid s_nC_{i,n})P(s_n)}{P(e_{i,n} \mid r_{i,n}C_{i,n})P(r_{i,n} \mid C_{i,n})P(C_{i,n})}, \tag{1}$$

其中，在(1)的右边，为了节省符号，我们分别用 $C_{i,n}$，$s_{i,n}$和 s_n 代替了 $N(C_{i,n})$，$N(s_{i,n})$和 $N(s_n)$，对于 $N(C_i)$等等，也是如此。整个这一节在证明中都遵从这个约定，但在陈述定理时并非如此。在简单学习中，由于公理 $E1$，$e_{i,n}$的条件概率在分子和分母中是

相同的,首先注意到这一点,我们可以简化(1)。其次,根据埃斯蒂斯和苏佩斯的文章(1959b)中证明的一般的学习定理,我们知道

$$P(r_{i,n+1} \mid e_{i,n}r_{i,n}s_{i,n}s_nC_{i,n}) = P(r_{i,n+1} \mid C'_{i,n+1}) = \frac{C_i + s - s_i}{N},$$

其中 $C_i' = C_i \cup (s - s_i)$。我们于是可能容易证明

$$P(r_{i,n} \mid s_{i,n}s_n) = \frac{s_i}{s}$$

和

$$P(s_{i,n} \mid s_nC_{i,n}) = \frac{\begin{pmatrix} C_i \\ s_i \end{pmatrix}\begin{pmatrix} n - C_i \\ s - s_i \end{pmatrix}}{\begin{pmatrix} N \\ s \end{pmatrix}}。$$

(在埃斯蒂斯和苏佩斯的文章(1959b)中,第一个等式的证明运用定理 2.4.1,第二个等式的证明运用定理 2.3.4。)

此外,如果这种抽样是空的,那么

$$P(r_{i,n+1} \mid e_{2,n}r_{i,n}s_n^OC_{i,n}) = P(r_{i,n} \mid C_{i,n})$$

对于这种空样本情况来说,这两项可删去;当然也有,$P(r_{i,n} \mid s_n^O) = \rho_i$。把这些不同的评论结合起来,并把它们应用于(1),我们推出 407

$$P(r_{i,n+1} \mid e_{i,n}r_{i,n}C_{i,n})$$

$$= \rho_iP(s^O) + \left[\sum_{s=s_i}^{N}\sum_{s_i=1}^{C_i}\frac{(C_i + s - s_i)}{N}\cdot\frac{s_i}{s}\cdot\frac{\begin{pmatrix} C_i \\ s_i \end{pmatrix}\begin{pmatrix} N - C_i \\ s - s_i \end{pmatrix}}{\begin{pmatrix} N \\ s \end{pmatrix}}P(s)\right]$$

$$\Big/ P(r_{i,n} \mid C_{i,n}), \tag{2}$$

算出(2)中 s_i 的和，运用这个定理之前提到的超几何分布的二阶原始矩的表达式和分布的期望值是$\frac{sC_i}{N}$的事实，我们有

$$
\begin{aligned}
&P(r_{i,n+1} \mid e_{i,n}r_{i,n}C_{i,n}) \\
&= \rho_i P(s^O)+\sum_{s=1}\left[\frac{C_i^2}{N^2}+\frac{sC_i}{N^2}-\frac{1}{sN^2}\left(sC_i+\frac{C_i(C_i-1)s(s-1)}{N-1}\right)\right]P(s) \\
&\Big/ P(r_{i,n} \mid C_{i,n}),
\end{aligned} \tag{3}
$$

现在对 s 求和，我们获得

$$
\begin{aligned}
&P(r_{i,n+1} \mid e_{i,n}r_{i,n}C_{i,n}) \\
&= \rho_i P(s^O)+(1-P(s^O))\left[\frac{C_i^2}{N^2}+\frac{\bar{s}C_i}{N^2}-\frac{C_i}{N^2}-\frac{C_i(C_i-1)(\bar{s}-1)}{N^2(N-1)}\right] \\
&\Big/ P(r_{i,n} \mid C_{i,n}),
\end{aligned} \tag{4}
$$

很容易表明，我们可以用$(1-P(s^O))\frac{C_i}{N}+\rho_i P(s^O)$取代 $P(r_{i,n} \mid C_{i,n})$。我们于是获得预期的结果。

证毕。

显然，如果我们假设，空样本的概率为零，这个定理就有一个更加简单的形式。

推论 1. 如果 $P(s^O)=0$，那么

$$
P(r_{i,n+1} \mid e_{i,n}r_{i,n}N(C_{i,n}))=\frac{(N-\bar{s})}{N-1}\frac{N(C_i)}{N}+\frac{\bar{s}-1}{N-1}。
$$

证明：基于这个推论的假设，我们根据这个定理立刻有

$$
\begin{aligned}
&P(r_{i,n+1} \mid e_{i,n}r_{i,n}C_i) \\
&= \frac{N}{C_i}\left[\frac{C_i^2}{N^2}+\frac{\bar{s}C_i}{N^2}-\frac{C_i}{N^2}-\frac{C_i(C_i-1)(\bar{s}-1)}{N^2(N-1)}\right]
\end{aligned}
$$

$$= \left(1 + \frac{1-\bar{s}}{N-1}\right)\frac{C_i}{N} + \frac{\bar{s}-1}{N} + \frac{\bar{s}-1}{N(N-1)}$$

$$= \frac{N-\bar{s}}{N-1}\frac{C_i}{N} + (\bar{s}-1)\left(\frac{1}{N-1}\right)$$

证毕。

第二个推论陈述了当 $N \to \infty$ 时的一个极限结果。我们所用的这个极限过程将在后面作出更详细的讨论，但在本质上，对于一个确定的实验者的程序表来说，我们选择刺激抽样模型$\mathfrak{S}(N)$的一个序列，使得对于每个 i，$\lim\limits_{N\to\infty}\frac{N(C_{i,1})}{N}$ 存在，$\lim\limits_{N\to\infty}\frac{\bar{s}(N)}{N}$ 存在，以及方差 $\frac{1}{N^2}\sum_{N(s)=0}^{N}(N(s)-\bar{s})^2P(s)$ 有一个零极限。(这样一个序列是能够被挑选出来的，稍后会使这一点更明显。)为了使与线性模型的联系在直观上更加明显，我们把线性模型的学习参数 θ 定义为

$$\theta = \lim_{N\to\infty}\frac{\bar{s}(N)}{N}。$$

我们于是有

推论 2: 408

$$\lim_{N\to\infty}P(r_{i,n+1} \mid e_{i,n}r_{i,n}N(C_{i,n})) = (1-\theta)\lim_{N\to\infty}\frac{C_i}{N} + \theta$$

下面四个定理及其推论在结构上类似于第一组，因此，在证明时省略了例行程序的细节。在第二定理中，我们需要的事实是：通过下列表达式，给出一个多变量的超几何分布的某个交叉矩(cross-moment)。

$$\sum_{a=0}^{r}\sum_{b=0}^{B} ab \frac{\binom{A}{a}\binom{B}{b}\binom{N-A-B}{s-a-b}}{\binom{N}{s}} = \frac{ABs(s-1)}{N(N-1)}$$

定理 2. 如果 $i \neq j$ 和 $P(e_{i,n}r_{j,n}N(C_{i,n})N(C_{j,n})) > 0$，那么

$$P(r_{i,n+1} | e_{i,n}r_{j,n}N(C_{i,n})N(C_{j,n}))$$

$$= \frac{\rho_j P(s^O)\left[(1-P(s^O))\dfrac{N(C_i)}{N} + \rho_i P(s^O)\right]}{(1-P(s^O))\dfrac{N(C_j)}{N} + \rho_j P(s^O)}$$

$$+(1-P(s^O))\left[\frac{N(C_i)N(C_j)}{N^2} + \frac{\bar{s}N(C_j)}{N^2} - \frac{N(C_i)N(C_j)(\bar{s}-1)}{N^2(N-1)}\right]$$

$$\Big/\left[(1-P(s^O))\frac{N(C_j)}{N} + \rho_j P(s^O)\right]。$$

证明：首先，我们有

$$P(r_{i,n+1} \mid e_{i,n}r_{j,n}C_{i,n}C_{j,n})$$

$$= \sum_{s=s_i+s_j}^{N}\sum_{s_i=0}^{C_i}\sum_{s_j=0}^{C_j} P(r_{i,n+1} \mid e_{i,n}r_{j,n}s_{i,n}s_{j,n}s_nC_{i,n})P(r_{j,n} \mid s_{j,n}s_n)$$

$$\cdot P(s_{i,n}s_{j,n} \mid s_nC_{i,n}C_{j,n})P(s_n)/P(r_{j,n} \mid C_{j,n})。\qquad (5)$$

在(5)的右边，我们已经从分子和分母中消去 $e_{i,n}$ 的条件概率。为了获得 $P(s_{i,n}s_{j,n} \mid s_nC_{i,n}C_{j,n})$，我们应用了埃斯蒂斯和苏佩斯(1959b)提出的定理 2.3.5。① 因此，通过与定理 1 所用的

① 这个定理是一个简单明了的抽样定理：

定理 2.3.5 如果 $P(N(s_n)C_n) > 0$，那么

$P(\bigcap_{i=1}^{r} N(s_{i,n}) \mid N(s_n)C_n) = \pi_{i=1}^{r}\binom{C_i}{s_i}\Big/\binom{N}{s}$，这里 r 是反应的个数。

论证一样的思路，我们从(5)推出

$$P(r_{i,n+1} \mid e_{i,n} r_{j,n} C_{i,n} C_{j,n})$$
$$= \rho_j P(s^O) P(r_{i,n} \mid C_{i,n}) / P(r_{j,n} \mid C_{j,n})$$
$$+ \left[\sum_{s=s_i+s_j}^{N} \sum_{s_i=0}^{C_i} \sum_{s_j=1}^{C_j} \cdot \frac{\binom{C_i}{s_i}\binom{C_j}{s_j}\binom{N-C_i-C_j}{s-s_i-s_j}}{\binom{N}{s}} P(s) \right]$$
$$\Big/ P(r_{j,n} \mid C_{j,n})。 \quad (6)$$

算出在(6)中的 s_i 与 s_j 的和，并运用在这个定理之前陈述的超几何分布的交叉矩，我们获得 409

$$P(r_{i,n+1} \mid e_{i,n} r_{j,n} C_{i,n} C_{j,n})$$
$$= \rho_j P(s^O) P(r_{i,n} \mid C_{i,n}) / P(r_{j,n} \mid C_{j,n})$$
$$+ \sum_{s=1}^{N} \left[\frac{C_i C_j}{N^2} + \frac{s C_j}{N^2} - \frac{1}{sN} \left[\frac{C_i C_j s(s-1)}{N(N-1)} \right] \right] P(s)$$
$$/ P(r_{j,n} \mid C_{j,n})。 \quad (7)$$

现在，对 s 求和，我们获得

$$P(r_{i,n+1} \mid e_{i,n} r_{j,n} C_{i,n} C_{j,n})$$
$$= \rho_j P(s^O) P(r_{i,n} \mid C_{i,n}) / P(r_{j,n} \mid C_{j,n})$$
$$+ (1 - P(s^O)) \left[\frac{C_i C_j}{N^2} + \frac{\bar{s} C_j}{N^2} - \frac{C_i C_j (\bar{s} - 1)}{N^2 (N-1)} \right]$$
$$\Big/ P(r_{j,n} \mid C_{j,n})。 \quad (8)$$

然后，用埃斯蒂斯和苏佩斯(1959b)的定理 2.4.5 代替(8)中的

$P(r_{i,n} \mid C_{i,n})$ 和 $P(r_{j,n} \mid C_{j,n})$，我们获得了预期的结果。①

证毕。

我们陈述两个推论，但没有证明。

推论 1. 如果 $P(s^O)=0$，那么

$$P(r_{i,n+1} \mid e_{i,n} r_{j,n} N(C_{i,n}) N(C_{j,n})) = \left(\frac{N-\bar{s}}{N-1}\right)\frac{N(C_i)}{N} + \frac{\bar{s}}{N}。$$

推论 2.

$$\lim_{N\to\infty} P(r_{i,n+1} \mid e_{i,n} r_{j,n} N(C_{i,n}) N(C_{j,n}))$$

$$= (1-\theta)\lim_{N\to\infty}\frac{N(C_i)}{N} + \theta。$$

我们现在考虑当试验 n 的强化不同于 e_i 时产生的两种情况。

定理 3. 如果 $k \neq i$，$k \neq 0$ 和 $P(e_{k,n} r_{i,n} N(C_{i,n}))>0$，那么

$$P(r_{i,n+1} \mid e_{k,n} r_{i,n} N(C_{i,n}))$$

$$= \rho_i P(s^O) + (1-P(s^O))\left[\frac{N(C_i)^2}{N^2} - \frac{N(C_i)}{N^2} - \frac{N(C_i)(N(C_i)-1)(\bar{s}-1)}{N^2(N-1)}\right]$$

$$\Big/\left[(1-P(s^O))\frac{N(C_i)}{N} + \rho_i P(s^O)\right]。$$

证明：与证明定理 1 的方式一样进行，但是，对不同强化求

---

① 已知对在一个给定试验中作出反应的条件性元素的个数，这个定理表达了一个直接的条件反应概率：

定理 2.4.5 如果 $P(C_{i,n})>0$，那么 $P(r_{i,n} \mid C_{i,n}) = (1-P(s_n^O))\frac{N(C_i)}{N} + P(s_n^O)\rho_{i,n}$。

和之后，调整了第一项，我们获得

$$P(r_{i,\,n+1} \mid e_{k,\,n}r_{i,\,n}N(C_{i,\,n}))$$

$$= \rho_i P(s^O) + \left[\sum_{s=i}^{N}\sum_{s_i=1}^{C_i}\frac{(C_i - s_i)}{N}\cdot\frac{s_i}{s}\frac{\binom{C_i}{s_i}\binom{N-C_i}{S-s_i}}{\binom{N}{s}}P(s)\right]$$

$$\Big/ P(r_{i,\,n} \mid C_{i,\,n})。\tag{9}$$

对 s_i 求和后得到 s，我们有 410

$$P(r_{i,\,n+1} \mid e_{k,\,n}r_{i,\,n}N(C_{i,\,n}))$$

$$= \rho_i P(s^O) + (1 - P(s^O))\left[\frac{C_i^2}{N^2} - \frac{C_i}{N^2} - \frac{C_i(C_i-1)(\bar{s}-1)}{N^2(N-1)}\right]$$

$$\Big/ P(r_{i,\,n} \mid C_{i,\,n}),\tag{10}$$

通过取代 $P(r_{i,\,n} \mid C_{i,\,n})$，从(10)中立刻获得预期的结果。

证毕。

推论 1. 如果 $P(s^O)=0$，那么

$$P(r_{i,\,n+1} \mid e_{k,\,n}r_{i,\,n}N(C_{i,\,n})) = \left(\frac{N-\bar{s}}{N-1}\right)\frac{N(C_i)}{N} - \frac{N-\bar{s}}{N(N-1)}。$$

推论 2.

$$\lim_{N\to\infty} P(r_{i,\,n+1} \mid e_{k,\,n}r_{i,\,n}N(C_{i,\,n})) = (1-\theta)\lim_{N\to\infty}\frac{N(C_i)}{N}。$$

下一个定理是对试验 n 和 $n+1$ 的不同反应。

定理 4. 如果 $k \neq i$，$k \neq 0$，$j \neq i$

和 $P(e_{k,\,n}r_{j,\,n}N(C_{i,\,n})N(C_{j,\,n})) > 0$，那么

$$P(r_{i,n+1} \mid e_{k,n}r_{j,n}N(C_{i,n})N(C_{j,n}))$$

$$=\frac{\rho_j P(s^O)\left[(1-P(s^O))\dfrac{N(C_i)}{N}+\rho_i P(s^O)\right]}{(1-P(s^O))\dfrac{N(C_j)}{N}+\rho_j P(s^O)}$$

$$+(1-P(s^O))\left[\frac{N(C_i)N(C_j)}{N^2}-\frac{N(C_i)N(C_j)(\bar{s}-1)}{N^2(N-1)}\right]$$

$$\Big/\left[(1-P(s^O))\frac{N(C_j)}{N}+\rho_j P(s^O)\right]。$$

证明：像定理 2 中的情形那样进行，但对不同的强化作出相应的改变，与证明定理 2 的(7)相对应，我们有

$$P(r_{i,n+1} \mid e_{i,n}r_{j,n}C_{i,n}C_{j,n})$$

$$=\rho_j P(s^O)P(r_{i,n} \mid C_{i,n})/P(r_{j,n} \mid C_{j,n})$$

$$+(1-P(s^O))\left[\frac{C_iC_j}{N^2}-\frac{C_iC_j(\bar{s}-1)}{N^2(N-1)}\right]\Big/P(r_{j,n} \mid C_{j,n})。$$

与前面一样，通过应用埃斯蒂斯和苏佩斯（1959b）的定理 2.4.5，很容易获得预期的结果。

证毕。

推论 1. 如果 $P(s^O)=0$，那么

$$P(r_{i,n+1} \mid e_{k,n}r_{j,n}N(C_{i,n})N(C_{j,n}))=\left(\frac{N-\bar{s}}{N-1}\right)\frac{N(C_i)}{N}。$$

推论 2.

$$\lim_{N\to\infty}P(r_{i,n+1} \mid e_{k,n}r_{j,n}N(C_{i,n})N(C_{j,n}))=(1-\theta)\lim_{N\to\infty}\frac{N(C_i)}{N}。$$

最后，我们有试验 n 的非强化定理。我们将注意到，无论 $j=i$ 还是 $j\neq i$，结果都是相同的。

定理 5. 如果 $P(e_{0,n}r_{j,n}N(C_{i,n})N(C_{j,n}))>0$，那么

$$P(r_{i,n+1} \mid e_{0,n}r_{j,n}N(C_{i,n})N(C_{j,n})) = P(r_{i,n} \mid N(C_{i,n}))。$$

证明：根据埃斯蒂斯和苏佩斯(1959b)的定理 2.5.6 和对 411 当前情形的明显扩展①，

$$\begin{aligned} P(r_{i,n+1} \mid e_{0,n}r_{j,n}N(C_{i,n})N(C_{j,n})) &= P(r_{i,n+1} \mid N(C_{i,n+1})) \\ &= P(r_i, n \mid N(C_{i,n}))。\end{aligned}$$

证毕。

包括序列 ω_n 的定理。与前一小节的定理相对应，我们现在有包括对序列 w_{n-1} 条件化的一个系列。② 为了更加简化表述，我们将假设，在所有这些定理中，$P(s^O)=0$。我们将不关心处理 $N\to\infty$ 的推论，因为我们将希望设 N 在一个稍微不同的阶段逼近无穷大。

定理 6. 如果 $P(s^O)=0$ 和 $P(e_{i,n}r_{i,n}\omega_{n-1})>0$，那么

$$P(r_{i,n+1} \mid e_{i,n}r_{i,n}\omega_{n-1})$$

$$= \frac{\frac{N-\bar{s}}{N-1}\sum_{N(C_i)=0}^{N} P(r_{i,n} \mid N(C_{i,n}))^2 P(N(C_{i,n}) \mid \omega_{n-1})}{P(r_{i,n} \mid \omega_{n-1})}$$

$$+\frac{\bar{s}-1}{N-1}。$$

---

① 所谈到的定理如下：

定理 2.5.6 如果 $P(e_{0,n}C_{i,n})>0$，那么 $P(C'_{i,n+1} \mid e_{0,n}C_{i,n}) = \begin{cases} 1 & \text{如果 } C'_i = C_i \\ 0 & \text{其他} \end{cases}$。

② 用一个更明显但较繁琐的等价类符号来表示，写成 $[\omega]_n$，而不是 ω_n，更适当，在这里，ω_n 只指实验 n。但为了避免这个附加符号，我总是写成 ω_n，而不是 $[\omega]_n$。注意这个形式符号定义如下：

$$[\omega]_n = \{\omega' : \omega'_i = \omega_i, \text{ for } 1 \leqslant i \leqslant n\}。$$

这个定义恰好在定义 1 之前就已经得到了非正式的陈述。

证明：我们根据 $N(C_i)$展开。(像前一节的证明情形那样，我们写 C_i，而不写 $N(C_i)$，等等。)

$$P(r_{i,n+1} \mid e_{i,n}r_{i,n}\omega_{n-1})$$

$$= \sum_{C_i=0}^{N} P(r_{i,n+1} \mid e_{i,n}r_{i,n}C_{i,n})P(r_{i,n} \mid C_{i,n})$$

$$\cdot P(C_{i,n} \mid \omega_{n-1})/P(r_{i,n} \mid \omega_{n-1})$$

$$= \sum_{C_i}\left[\frac{(N-\bar{s})}{(N-1)}\frac{C_i}{N} + \frac{\bar{s}-1}{N-1}\right]P(r_{i,n} \mid C_{i,n})P(C_{i,n} \mid \omega_{n-1})$$

$$/P(r_{i,n} \mid \omega_{n-1})$$

$$= \left[\frac{N-\bar{s}}{N-1}\sum_{C_i}P(r_{i,n} \mid C_{i,n})^2 P(C_{i,n} \mid \omega_{n-1})\right.$$

$$\left. + \frac{\bar{s}-1}{N-1}P(r_{i,n} \mid \omega_{n-1})\right]\Big/P(r_{i,n} \mid \omega_{n-1}),$$

现在，我们通过在右边第二项的分子和分母中约去 $P(r_{i,n} \mid \omega_{n-1})$，获得结论。

证毕。

注意，由于相关的强化公理和已知整个序列 ω_{n-1} 的事实，所以，在第一行右边的分子和分母中约去 $e_{i,n}$的条件概率。

下面四个定理的证明与此类似，其中，每一个定理都依赖于前一小节相对应的定理和推论，因此，我们省略了多数定理的证明。

412 **定理 7.** *如果* $P(s^O) = 0$，$i \neq j$ *和* $P(e_{i,n}r_{j,n}\omega_{n-1}) > 0$，*那么*

$$P(r_{i,n+1} \mid e_{i,n}r_{j,n}\omega_{n-1})$$

$$= \frac{N-\bar{s}}{N-1}\sum_{N(C_i)=0}^{N}\sum_{N(C_j)=0}^{N-N(C_i)}P(r_{i,n} \mid N(C_{i,n}))$$

$$\cdot P(r_{j,n} \mid N(C_{j,n}))P(N(C_{i,n})N(C_{j,n}) \mid \omega_{n-1})$$

$$/P(r_{j,n} \mid \omega_{n-1}) + \frac{\bar{s}}{N}。$$

证明：我们根据 $N(C_{i,n})N(C_{j,n})$ 展开，并且运用定理 2。

$$P(r_{i,n+1} \mid e_{i,n}r_{j,n}\omega_{n-1})$$

$$= \sum_{C_i=0}^{N} \sum_{C_j=0}^{N-C_i} P(r_{i,n+1} \mid e_{i,n}r_{j,n}C_{i,n})P(r_{j,n} \mid C_{j,n})$$

$$\cdot P(C_{i,n}C_{j,n} \mid \omega_{n-1})/P(r_{j,n} \mid \omega_{n-1})$$

$$= \sum_{C_i} \sum_{C_j} \left[\frac{(N-\bar{s})}{(N-1)}\frac{C_i}{N} + \frac{\bar{s}}{N}\right]\frac{C_j}{N}P(C_{i,n}C_{j,n} \mid \omega_{n-1})$$

$$/P(r_{j,n} \mid \omega_{n-1})$$

$$= \frac{N-\bar{s}}{N-1}\sum_{C_i}\sum_{C_j}P(r_{i,n} \mid C_{i,n})P(r_{j,n} \mid C_{j,n})$$

$$\cdot P(C_{i,n}C_{j,n} \mid \omega_{n-1})/P(r_{j,n} \mid \omega_{n-1}) + \frac{\bar{s}}{N},$$

因为

$$\frac{\bar{s}}{N}\sum_{C_i}\sum_{C_j}\frac{C_j}{N}P(C_{i,n}C_{j,n} \mid \omega_{n-1})/P(r_{j,n} \mid \omega_{n-1})$$

$$= \frac{\bar{s}}{N}\sum_{C_j}P(r_{j,n} \mid C_{j,n})$$

$$\cdot P(C_{j,n} \mid \omega_{n-1})/P(r_{j,n} \mid \omega_{n-1})$$

$$= \frac{\bar{s}}{N}\frac{P(r_{j,n} \mid \omega_{n-1})}{P(r_{j,n} \mid \omega_{n-1})} = \frac{\bar{s}}{N}。$$

证毕。

定理 8. 如果 $P(s^O)=0$，$k \neq i$，$k \neq 0$ 和 $P(e_{k,n}r_{i,n}\omega_{n-1}) >$

0,那么

$$P(r_{i,\,n+1} \mid e_{k,\,n} r_{i,\,n} \omega_{n-1})$$

$$= \frac{\frac{N-\bar{s}}{N-1} \sum_{N(C_i)=0}^{N} P(r_{i},\, n \mid N(C_i,\, n))^2 P(N(C_i,\, n) \mid w_{n-1})}{P(r_i,\, n \mid w_{n-1})}$$

$$+ \frac{N-\bar{s}}{N(N-1)}。$$

定理 9. 如果 $P(s^O)=0$, $k \neq i$, $k \neq 0$, $j \neq i$ 和 $P(e_{k,\,n} r_{j,\,n} \omega_{n-1})>0$,那么

$$P(r_{i,\,n+1} \mid e_{k,\,n} r_{j,\,n} \omega_{n-1})$$

$$= \frac{N-\bar{s}}{N-1} \sum_{N(C_i)=0}^{N} \sum_{N(C_j)=0}^{N-N(C_i)} P(r_{i,\,n} \mid N(C_{i,\,n})) P(r_{j,\,n} \mid N(C_{j,\,n}))$$

$$\cdot P(N(C_{i,\,n}) N(C_{j,\,n}) \mid \omega_{n-1}) / P(r_{j,\,n} \mid \omega_{n-1})。$$

与定理 5 相对应,我们现在需要依赖于 $j=i$ 或者 $j \neq i$ 的两个定理。

定理 10. 如果 $P(s^O)=0$ 和 $P(e_{0,\,n} r_{i,\,n} \omega_{n-1})>0$,那么

$$P(r_{i,\,n+1} \mid e_{0,\,n} r_{i,\,n} \omega_{n-1})$$

$$= \sum_{N(C_i)=0}^{N} P(r_{i,\,n} \mid N(C_{i,\,n}))^2 P(N(C_{i,\,n}) \mid \omega_{n-1}) / P(r_{i,\,n} \mid \omega_{n-1})。$$

413 **定理 11.** 如果 $P(s^O)=0$, $j \neq i$ 和 $P(e_{0,\,n} r_{j,\,n} \omega_{n-1})>0$,那么

$$P(r_{i,\,n+1} \mid e_{0,\,n} r_{j,\,n} \omega_{n-1})$$

$$= \sum_{N(C_i)=0}^{N} \sum_{N(C_j)=0}^{N-N(C_i)} P(r_{i,\,n} \mid N(C_{i,\,n})) P(r_{j,\,n} \mid N(C_{j,\,n}))$$

$$\cdot P(N(C_i) N(C_j) \mid \omega_{n-1}) / P(r_{j,\,n} \mid \omega_{n-1})。$$

正如刚才所陈述的六个定理所建议的那样,下一步是证明

当 $N \to \infty$ 时，

$$\sum_{N(C_i)=0}^{N} P(r_{i,n} | N(C_{i,n}))^2 P(N(C_{i,n}) \mid \omega_{n-1}) = P(r_{i,n} | \omega_{n-1})^2 \tag{11}$$

和

$$\sum_{N(C_i)=0}^{N} \sum_{N(C_j)=0}^{N-N(C_i)} P(r_{i,n} \mid N(C_{i,n})) P(r_{j,n} \mid N(C_{j,n})) P(N(C_{i,n}) N(C_{j,n}) \mid \omega_{n-1}) = P(r_{i,n} \mid \omega_{n-1}) P(r_{j,n} \mid \omega_{n-1}) \tag{12}$$

其中 $i \neq j$。为达到这个目标，方便的是，定义某些方差和协方差。

$$\sigma_{i\omega}^2(N, n) = \sum_{N(C_i)=0}^{N} [P(r_{i,n} \mid N(C_{i,n})\omega_{n-1}) - P(r_{i,n} \mid \omega_{n-1})]^2 \cdot P(N(C_{i,n}) \mid \omega_{n-1}), \tag{13}$$

$$\sigma_{ij\omega}(N, n) = \sum_{N(C_i)=0}^{N} \sum_{N(C_j)=0}^{N-N(C_i)} [P(r_{i,n} \mid N(C_{i,n})\omega_{n-1}) - P(r_{i,n} \mid \omega_{n-1})] \cdot [P(r_{j,n} \mid N(C_{j,n})\omega_{n-1}) - P(r_{j,n} \mid \omega_{n-1})] P(N(C_{i,n}) N(C_{j,n}) \mid \omega_{n-1})。 \tag{14}$$

容易表明

$$\sigma_{i\omega}^2(N, n) = \sum_{N(C_i)=0}^{N} P(r_{i,n} | N(C_{i,n}))^2 P(N(C_{i,n}) | \omega_{n-1}) - P(r_{i,n} | \omega_{n-1})^2 \tag{15}$$

和

$$\sigma_{ij\omega}(N, n) = \sum_{N(C_i)=0}^{N} \sum_{N(C_j)=0}^{N-N(C_i)} P(r_{i,n} \mid N(C_{i,n}))P(r_{j,n} \mid N(C_{j,n}))$$

$$\cdot P(N(C_{i,n})N(C_{j,n}) \mid \omega_{n-1})$$

$$- P(r_{i,n} \mid \omega_{n-1})P(r_{j,n} \mid \omega_{n-1})。 \tag{16}$$

从(15)和(16)显然看出,如果

$$\lim_{N\to\infty}\sigma_{i\omega}^2(N, n) = 0 \tag{17}$$

和

$$\lim_{N\to\infty}\sigma_{ij\omega}(N, n) = 0, \tag{18}$$

那么,倘若(15)和(16)右边的个别项的极限存在,将确立(11)和(12)。因此,这种存在性确实是从我们的其他结果得出的结论。在我们继续确立(17)和(18)之前,我们重新考虑当 $N\to\infty$时的极限假设。

414 **极限假设**。对于一个确定的实验者的程序表来说,因为呈现集合是不变的,这恰好是一个确定的强化程序表,而且,已知 $r_{i,n}$,对于强化 $e_{k,n}$的固定概率来说,我们选择刺激-抽样模型的一个序列($\mathfrak{S}(1)$, $\mathfrak{S}(2)$, …, $\mathfrak{S}(N)$, …),其中,这个序列的第 N 项有 N 个刺激元素并满足下列三个条件:

Ⅰ. 对于每个 N,存在数 $\gamma_i^{(N)} \geqslant 0$,其中,$\sum\gamma_i^{(N)} = 1$,使得对于试验 1 中的任意条件作用函数 $C_i^{(N)}$

$$P_N(C_i^{(N)}) = 0$$

除非 $\dfrac{N(C_i^{(N)})}{N} = \gamma_i^{(N)}$,并且存在数 $\gamma_i \geqslant 0$,其中,$\sum\gamma_i = 1$,使得

$$\lim_{N\to\infty}\gamma_i^{(N)} = \gamma_i。$$

Ⅱ. 设$\bar{s}(N)$是$\mathfrak{S}(N)$的预期的样本大小(在每次试验中),于是,存在$\lim\limits_{N\to\infty}\frac{\bar{s}(N)}{N}$,并且,我们把$\theta$定义为这个极限。

Ⅲ. 当$N\to\infty$,$\frac{N(s)}{N}$的分布方差趋于0,即

$$\lim_{N\to\infty}\frac{1}{N^2}\sum_{N(s)=0}^{N}(N(s)-\bar{s})^2P_N(s)=0。$$

因为这三个规定只把这些限制强加于这些模型的试验1的条件作用函数和与n无关的抽样分布,所以,显然有可能为任意数$\gamma_i\geqslant 0\left(其中,\sum\gamma_i=1\right)$和$\theta$(其中,$0<\theta\leqslant 1$)选择这样一个序列。(Ⅱ)和(Ⅲ)的直观内容似乎是明晰的,但(Ⅰ)有些令人惊讶。为了把一个特定线性模型表示为刺激-抽样模型的一个序列的极限,有必要确定与试验1的每个反应相关的每种刺激的极限比。在作出这个最后的陈述时,我想到了前面给出的线性模型的描述。我后面会再回到这一点。应该明显的是,可以用几种无关紧要的方式弱化(Ⅰ)。

有必要一起确定(17)和(18),因为我们是通过归纳进行的,并且,对于某些情况来说,对(17)或(18)的$n+1$的论证依赖于能假设另一个极限对于n成立。预期的表证定理是这个定理的直接推论。

定理12. 如果$P(\omega_n)>0$,那么

(i) $$\lim_{N\to\infty}\sigma_{i\omega}^2(N,n)=0$$

以及

(ii) $$\lim_{N\to\infty}\sigma_{ij\omega}(N,n)=0。$$

证明: 正如已经提到的那样,我们通过归纳从这个定理的两部分同时进行。

对于 $n = 1$，这些结果直接来自选择规则(Ⅰ)。例如，如果 $\frac{N(C_{i,1})}{N} = \gamma_i^{(N)}$，那么 $P(N(C_{i,1})) = 1$ 和

$$P(r_{i,1} \mid N(C_{i,1}))^2 P(N(C_{i,1})) = P(r_{i,1})^2。$$

415 现在假设，这个定理对于 n 成立，我们必须考虑几种情况。[通常，在下面证明时，我们写 C_i，而不写 $N(C_i)$，等等。]

情形 1. $\omega_n = e_{k,n} r_{i,n} \omega_{n-1}$。

对于这种情况，(i)的论证不依赖于(ii)，但情形 2 并非如此。我们首先确定(i)。

$$\sum_{C_i} P(r_{i,n+1} \mid C_{i,n+1})^2 P(C_{i,n+1} \mid \omega_n)$$

$$= \sum_{C_i} \sum_{C_i'} \sum_{s=s_i}^{N} \sum_{s_i=0}^{C_i'} P(r_{i,n+1} \mid C_{i,n+1})^2 P(C_{i,n+1} \mid e_{k,n} r_{i,n} s_{i,n} s_n C'_{i,n})$$

$$\cdot P(r_{i,n} \mid s_{i,n} s_n) P(s_{i,n} \mid s_n C_{i,n}) P(s_n) P(C'_{i,n} \mid \omega_{n-1})$$

$$/ P(r_{i,n} \mid \omega_{n-1})。 \tag{19}$$

现在根据前面的结果，我们知道，$C_{i,n+1}$ 的条件概率为零，除非 $C_{i,n+1} = C_i' - s_i$，据此，我们作出这种简化，排除了对 C_i 的求和，并且，用熟悉的表达式代替(19)右边的其他项得到

$$\sum_{C_i} P(r_{i,n+1} \mid C_{i,n+1})^2 P(C_{i,n+1} \mid \omega_n)$$

$$= \sum_{C_i'} \sum_{s} \sum_{s_i} \left(\frac{C_i' - s_i}{N}\right)^2 \frac{s_i}{s} P(s_{i,n} \mid s_n C_{i,n}) P(s_n) P(C'_{i,n} \mid \omega_{n-1})$$

$$/ P(r_{i,n} \mid \omega_{n-1})。 \tag{20}$$

当 $N \to \infty$ 时，$\frac{s_i}{s} \to \frac{C_i}{N}$ 和 $\frac{\bar{s}}{N} \to \theta$，并且，抽样分布的方差趋

于零。① 因此,在极限(20)中约简为

$$\lim_{N\to\infty}\sum_{C_i}P(r_{i,\,n+1}\mid C_{i,\,n+1})^2P(C_{i,\,n+1}\mid\omega_n)$$

$$=\lim\sum_{C_i}(1-\theta)^2\left(\frac{C_i}{N}\right)^3\cdot P(C_{i,\,n}\mid\omega_{n-1})/P(r_{i,\,n}\mid\omega_{n-1})$$

$$=\lim(1-\theta)^2\sum_{C_i}P(r_{i,\,n}\mid C_{i,\,n})^3\cdot P(C_{i,\,n}\mid\omega_{n-1})$$

$$/P(r_{i,\,n}\mid\omega_{n-1})。\tag{21}$$

我们现在利用我们的归纳假设。我们首先注意到,(21)的右边是用 $P(r_{i,\,n}|C_{i,\,n})$的试验 n 的三阶原始矩 α_3 来表示的,这可以根据下列等式用三阶中心矩 μ_3、二阶原始矩 α_2 和平均值 μ 来表示:

$$\alpha_3=\mu_3+3\alpha_2\mu-2\mu^3,$$

但是,根据我们的归纳假设,所有的中心矩都为零,因此,$\mu_3=0$ 和 $\alpha_2=\mu^2$,据此,

$$\alpha_3=\mu^3=\lim P(r_{i,\,n}\mid\omega_{n-1})^3,\tag{22}$$

合并(21)和(22),我们有 416

$$\lim\sum_{C_i}P(r_{i,\,n+1}\mid C_{i,\,n+1})^2P(C_{i,\,n+1}\mid\omega_n)$$

$$=(1-\theta)^2\lim P(r_{i,\,n}\mid\omega_{n-1})^2,\tag{23}$$

我们现在根据我们的归纳假设 $\alpha_2=\mu^2$ 的事实应用于当 $N\to\infty$ 时的定理 8,并获得

---

① 可能会注意到,在这一点,我们使得下列假设的运用成为至关重要的:即把抽样方差 $\frac{1}{N^2}\sum_{N(s)=0}^{N}(N(s)-\bar{s})^2P(s)$ 趋于零作为一个极限,因为如果这个极限不为零,需要在这个极限中明确计算 s_i 的超几何分布及 s 的抽样分布的各种原始距。

$$\lim P(r_{i,\,n+1} \mid \omega_n) = (1-\theta)\lim P(r_{i,\,n} \mid \omega_{n-1})。\quad (24)$$

对于情形 1 来说,等式(i)是直接从(23)和(24)得出的。

我们现在转向(ii),运用使我们从(19)推出(20)的同样论证。

$$\begin{aligned}
&\alpha(i,\, j,\, \omega,\, n+1)\\
&= \sum_{C_i}\sum_{C_j} P(r_{i,\,n+1} \mid C_{i,\,n+1}) P(r_{j,\,n+1} \mid C_{j,\,n+1})\\
&\quad \cdot P(C_{i,\,n+1}C_{j,\,n+1} \mid \omega_n)\\
&= \sum_{C_i}\sum_{C_j}\sum_{s}\sum_{s_i}\sum_{s_j} \frac{(C_i - s_i)}{N}\frac{(C_j - s_j)}{N}\frac{s_i}{s} P(s_{i,\,n}s_{j,\,n} \mid s_n C_{i,\,n}C_{j,\,n})\\
&\quad \cdot P(s_n)P(C_{i,\,n}C_{j,\,n} \mid \omega_{n-1})/P(r_{i,\,n} \mid \omega_{n-1})。\quad (25)
\end{aligned}$$

现在,再一次,当 $N \to \infty$ 时,$\frac{s_i}{s} \to \frac{C_i}{N}$,$\frac{s_j}{s} \to \frac{C_j}{N}$ 和 $\frac{\bar{s}}{N} \to \theta$,以及 $\frac{1}{N_2}\sum_s (s-\bar{s})^2 P(s) \to 0$,据此,在极限(25)中简化为

$$\begin{aligned}
&\lim \alpha(i,\, j,\, \omega,\, n+1)\\
&= (1-\theta)^2 \lim \sum_{C_i}\sum_{C_j} \left(\frac{C_i}{N}\right)^2 \left(\frac{C_j}{N}\right) P(C_{i,\,n}C_{j,\,n} \mid \omega_{n-1})\\
&\quad / P(r_{i,\,n} \mid \omega_{n-1})\\
&= (1-\theta)^2 \lim \sum_{C_i}\sum_{C_j} P(r_{i,\,n} \mid C_{i,\,n})^2 P(r_{j,\,n} \mid C_{j,\,n})\\
&\quad \cdot P(C_{i,\,n}C_{j,\,n} \mid \omega_{n-1})/P(r_{i,\,n} \mid \omega_{n-1})。\quad (26)
\end{aligned}$$

我们现在运用我们的归纳假设计算试验 n 的原始交叉矩,这通过(26)右边的表达式来表示。因此,不再提符号 ω 和 n,设

$$\alpha(i^2,\, j) = \sum_{C_i}\sum_{C_j} P(r_{i,\,n} \mid C_{i,\,n})^2 P(r_{j,\,n} \mid C_{j,\,n}) P(C_{i,\,n}C_{j,\,n} \mid \omega_{n-1}),\quad (27)$$

$$\mu(i^2, j) = \sum_{C_i} \sum_{C_j} [P(r_{i,n} | C_{i,n}) - \mu(i)]^2 [P(r_{j,n} | C_{i,n}) - \mu(j)] \cdot P(C_{i,n} C_{j,n} \mid \omega_{n-1}) \tag{28}$$

这里,$\mu(i)$是 $P(r_{i,n} | C_{i,n})$的平均值,$\mu(j)$也一样。而且,$\alpha(i, j)$的定义应该是明显的。现在,通过明确的计算

$$\mu(i^2, j) = \alpha(i^2, j) - \alpha(i^2)\mu(j) - 2\alpha(i, j)\mu(i) + 2\mu(i)^2\mu(j) + \mu(i)^2\mu(j) - \mu(i)^2\mu(j)。\tag{29}$$

从我们的归纳容易得出,中心交叉矩 $\mu(i^2, j) = 0$。也根据我们的归纳假设,得出 $\mu(i^2) = \sigma_i^2 = 0$,据此,

$$\alpha(i^2) = \mu(i)^2。$$

还有 417

$$\mu(i, j) = \sigma_{ij} = \alpha(i, j) - \mu(i)\mu(j),$$

但再次根据我们的归纳假设,$\sigma_{ij} = 0$,据此

$$\alpha(i, j) = \mu(i)\mu(j), \tag{30}$$

将这些结论代入(29),我们得出结论

$$\alpha(i^2, j) = \mu(i)^2\mu(j)。\tag{31}$$

把(31)应用于(26),我们直接得到

$$\lim \alpha(i, j, \omega, n+1) = (1-\theta)^2 \lim P(r_{i,n} | \omega_{n-1}) P(r_{j,n} | \omega_{n-1})。\tag{32}$$

现在正如对(i)的论证那样,根据我们的归纳假设,$\alpha(i) = \mu(i)^2$ 和(24)成立。此外,根据(30)和定理 4,对于 $\omega_n = e_{k,n} r_{i,n} \omega_{n-1}$,

$$\lim P(r_{j,n+1} \mid \omega_n) = (1-\theta) P(r_{j,n} \mid \omega_{n-1}), \tag{33}$$

对于情形 1 来说,等式(ii)直接从(24),(32)和(33)得出。

我们将不详细考虑所有的其他情况，但我们将再概述几种情况的主要论证。情形 2 的有趣之处是，它解释了为什么对 σ_i^2 和 σ_{ij} 必须同时进行归纳的原因。

情形 2. $\omega_n = e_{k,\,n} r_{j,\,n} \omega_{n-1}$。

与前面一样，我首先考虑(i)。根据情形 1 所用的方法，但适合于 $j \neq i$ 的情况，我们有

$$\begin{aligned}\alpha(i^2,\ n+1) &= \sum_{C_i} P(r_{i,\,n+1} \mid C_{i,\,n+1})^2 P(C_{i,\,n+1} \mid \omega_n) \\ &= \sum_{C_i}\sum_{C_j}\sum_{s}\sum_{s_i}\sum_{s_j}\left(\frac{C_i - s_i}{N}\right)^2 \frac{s_j}{s} P(s_i s_j \\ &\quad \mid s_n C_{i,\,n} C_{j,\,n}) P(s_n) P(C_{i,\,n} C_{j,\,n} \mid \omega_{n-1}) \\ &\quad / P(r_{j,\,n} \mid \omega_{n-1})。\end{aligned} \tag{34}$$

因此，当 $N \to \infty$ 时，

$$\begin{aligned}&\lim \alpha(i^2,\ n+1) \\ &= (1-\theta)^2 \lim \sum_{C_i}\sum_{C_j}\left(\frac{C_i}{N}\right)^2 \frac{C_j}{N} P(C_{i,\,n} C_{j,\,n} \mid \omega_{n-1}) \\ &\quad / P(r_{j,\,n} \mid \omega_{n-1}) \\ &= (1-\theta)^2 \lim \sum_{C_i}\sum_{C_j} P(r_{i,\,n} \mid C_{i,\,n})^2 P(r_{j,\,n} \mid C_{j,\,n}) \\ &\quad \cdot P(C_{i,\,n} C_{j,\,n} \mid \omega_{n-1}) / P(r_{j,\,n} \mid \omega_{n-1})。\end{aligned} \tag{35}$$

(35)右边的表达式与(26)右边的表达式相同，而且，我们正好应用了同样的方法，在过程中运用对 σ_{ij} 的归纳假设得到

$$\lim \alpha(i^2,\ n+1) = (1-\theta)^2 \lim P(r_{i,\,n} \mid \omega_{n-2})^2, \tag{36}$$

与前面一样，从(36)和定理 4，很容易获得等式(i)。

418 现在转向(ii)，根据通常的方法，我们有

$$\alpha(i, j, n+1)$$

$$= \sum_{C_i} \sum_{C_j} \sum_{s} \sum_{s_i} \sum_{s_j} \frac{(C_i - s_i)}{N} \frac{(C_j - s_j)}{N} \frac{s_j}{s}$$

$$\cdot P(s_{i,n} S_{j,n} \mid s_n C_{i,n} C_{j,n}) P(s_n) P(C_{i,n} C_{j,n} \mid \omega_{n-1})$$

$$/P(r_{i,n} \mid \omega_{n-1})。 \tag{37}$$

但(37)的右边正好与 i 与 j 交换后的(25)的右边一样，以同样的方法进行论证。

情形 3. $\omega_n = e_{k,n} r_{j',n} \omega_{n-1}$。

(i)的证明与用 j' 取代 j 的情形 2 一样。这种情况不同于关于(ii)的证明，因为必须考虑一种新的原始交叉矩，即 $\alpha(i, j, j')$，然而，可以与 $\alpha(i^2, j)$ 一样的方式来对待这个问题。

$$\alpha(i, j, j', n+1)$$

$$= \sum_{C_i} \sum_{C_j} \sum_{s} \sum_{s_i} \sum_{s_j} \sum_{s_j'} \frac{(C_i - s_i)}{N} \frac{(C_j - s_j)}{N} \frac{s_j'}{s}$$

$$\cdot P(s_{i,n} s_{j,n} s_{j',n} \mid s_n C_{i,n} C_{j,n} C_{j',n}) P(s_n) P(C_{i,n} C_{j,n} C_{j',n} \mid \omega_{n-1})$$

$$/P(r_{j',n} \mid \omega_{n-1})。 \tag{38}$$

因此，当 $N \to \infty$ 时，

$$\lim \alpha(i, j, j', n+1) = (1-\theta)^2 \lim \alpha(i, j, j', n)/\mu(j', n), \tag{39}$$

但是中心交叉矩(在试验 n 中)是

$$\mu(i, j, j') = \alpha(i, j, j') - \alpha(i, j)\mu(j') - \alpha(i, j')\mu(j)$$
$$- \alpha(j, j')\mu(i) - 2\mu(i)\mu(j)\mu(j')。 \tag{40}$$

根据我们的归纳，假设 $\mu(i, j, j') = 0$ 和 $\alpha(i, j) = \mu(i)\mu(j)$，等等，据此

$$\alpha(i,j,j') = \mu(i)\mu(j)\mu(j')。 \tag{41}$$

正如前面的结果所预期的那样。从(39)和(41)我们有

$$\lim \alpha(i,j,j',n+1) = (1-\theta)^2 \lim \mu(i,n)\,\mu(j,n), \tag{42}$$

其余的论证按照标准方法继续进行。

以类似的方式证明其他六种情形,因而省略。应该明显的是,这种论证总是取决于确立像(30)、(31)或(41)那样的一个结果。

证毕。

作为本节中的定理的一个结果,我们现在有下列线性模型表征定理。

定理 13. (表征定理)已知任意单一参数线性模型$(R, E, \Omega, P, \theta)$,存在刺激-抽样模型的一个序列$(\mathfrak{S}(1), \cdots, \mathfrak{S}(N), \cdots)$,使得对于每一个$\omega$和$n$,

$$P(\omega_n) = \lim_{N\to\infty} P_N(\omega_n)$$

这里P_N是$\mathfrak{S}(N)$的概率测度。

证明: 我们用选择规则(Ⅰ)、(Ⅱ)和(Ⅲ),即本节前面的定理,并继续对n进行归纳。对于$n=1$,我们运用选择规则(Ⅰ)。假设这个定理对n成立,我们于是运用了这样的事实:在极限状态下这些公理对于刺激-抽样模型的序列成立。其余的简单归纳证明正好类似于埃斯蒂斯和苏佩斯关于线性模型(1959a)的定理 4.1 的证明,这里将不再重复。

419

正如从这两节所参考的主要的文章资料能够看出的那样,这些主要结果是三十多年前获得的。我仍然喜欢把它们作为数学心理学中表征定理的例子。根据某些工作,也许能简化定理 13 的证明,但即使这样,对于为即便是相对简单的科学理论提供准确的还原定理的技术问题来说,它仍然将作为证言。由于

有可能在这个定理及其证明中作出一些改变,所以,一个新的应用领域可能是,现在实验经济学中用到的强化学习模型。至少在过去十年内,经济学家比心理学家对表征定理更感兴趣。

8.5 十种语言的理解性文法的机器人学习[①]

我没有研究所有的细节,只希望根据机器人听到一句话之后所发生的各种事件,来表达学习自然语言过程的一种相当清楚的直观感觉。(但应该知道,机器学习的基本程序,如果没有重大修改,也适合于其他应用,例如,物理学中的应用题的机器学习,这是我后面扼要考虑的问题。)同样,在下列标准学习用法中,我将经常谈到试验,当然在这种情况下,我的意思是说,试验是从命令机器人执行一句话开始的。

从概念上描述该程序体现的学习过程的最重要的方式是,在每个试验一开始时,描述一个机器人的记忆状态。由于学习,这种记忆在四个方面发生变化。首先是在一种给定语言的语词和机器人世界中作为指示行动、对象、特性和关系的内部符号之间的关联关系。一个中心问题是,学习在每一种语言中与内部符号真正相关的词是什么。这种记忆的第二个方面是,一个给定语词的指示价值发生了变化,这将影响到与它相关联的概率。第三个方面是,保存在试验期间有效的一个特定口令的短期记忆发生了变化。在给出一个特殊命令的试验之后,这种记忆内

① 本节基于苏佩斯、伯特纳(Böttner)和良(Liang)的文章(1996)以及苏佩斯和良的文章(1996)。之前,最初在1991年提交的两篇会议论文,描述了这条进路,这两篇文章是,苏佩斯、伯特纳和良的文章(1992)以及苏佩斯、伯特纳和良的文章(1995)。

容发生衰减，而且，不适用于存取。这意味着，试验一开始，在发出命令之前，这个短期缓冲器是空的。迄今为止由于某种扩展，我所说的内容符合经典的联想理论，但对于语言学习来说，很显然，这种受语义限制的联想关系和短期记忆的某些简单特征肯定是很不够的。

420 第四个方面是学习文法形式的一个重要方面。考虑口令，*Get the nut*（接住这个果仁）。这是文法形式 *A the O* 的一个例子，这里，*A* 是行动范畴，*O* 是对象范畴。这种形式实际上代表了有点过分简单化的形式，因为我们不仅仅有行动的单一范畴。存在着几个子范畴，这些范畴依赖于所需要的自变量的个数，还有某些其他自然的语义要求。然而，这个例子举例说明了操作原理。只对向机器人发出口令的实际例子的概括，得出这种文法形式。被学习的自然语言的任何一类文法的先前知识，对于机器人来说，都是没有的。同样重要的事实是，与被认为是概括出来的每一种文法形式相关的是，成为概括基础的词的联想及其内部表征。例如，如果 *Get the nut* 是生成刚才陈述的文法形式发生的事，那么联想 *get*～$ *g* 和 *nut*～$ *n* 也与那种文法形式存储在一起，这里，$ *g* 和 $ *n* 是机器人知道其所指的内部符号。当通过进一步的学习删除了不正确的联想时，也就删除了基于这种联想的文法形式。

理解性文法。大多数的语言学分析关注详细到足以产生所研究语言的自然表达的文法。相反，正如我们这里所描述的，一种理解性文法能生成表达的一个超集。要求这些规则只为通向一种语言表达的正确的语义解释。机器人，像幼儿一样，在它们能说话之前，很容易会有能力理解语言。虽然收集关于幼儿生成的理解性语法的准确而完整的资料，是困难的和微妙的，但是，证据是颠覆性的，机器人的理解力远大于它们的说话能力。

指示问题。在我们一直阐述的机器学习自然语言的概率理论中,我们在对自然语言进行语义分析时,已经遇到了一个新形式的标准问题,即如何处理非指示词。我们不是在某种绝对意义上,而是在相对于一个语义范畴的固定集合的意义上,意指非指示。大致说来,在机器人情况下,这些范畴是行动、对象、特性和关系的范畴。很可能在某些详尽阐述的自然语言的集合论语义学中,像定冠词 *the* 之类的非指示词,表示一个复杂的集合论函数,但是,这样一种详尽阐述的语义与语言学习的相关性是令人怀疑的。在机器人的语境中,我们有一些概念更简单和更接近于普通人关于指示是什么的观点。我们把颜色和对象词、普通名词、熟悉的具体行动的词,等等,作为指示词。我们在英语中用普通介词,有时在其他语言中用其他术语,表示大多数情形中的关系。另一方面,我们根据物理量的内部方程式语言,进一步从常识中删除了以物理应用题为核心的计算的语义。

当学习第一语言的一个小孩子,或者学习第二语言的一位老人,第一次遇到那种新语言中的表达时,标出非指示词的方式是完全不同的。有证据表明,在英语和其他语言中,运用各种超音质特征,对孩子是有帮助的。例如,在为幼儿提供的许多表达中,并不强调定冠词或不定冠词,而是调整普通名词,像在 *Hand me the cup* 的表达中那样。但是,这种方式似乎并不统一,无论如何,并不自然地适用于我们对机器学习的研究,在机器学习的情况下,我们使用书面输入文字,没有附加超音质符号。 421

正如已经明确的那样,我们的机器学习进路的一个核心特征是,在被学习的自然语言的词和内部语言的指示符号之间的概率的联想。恰当的是,在开始时同等地对待所有的词,因此,基于一种均匀分布的抽样形成联想。另一方面,在机器人已经

学习了许多词和获得了大量语言之后，如果向机器人发出一个难理解的命令，例如，*Get the astrolabe*，让内部符号 *$ ast* 以相等的概率与定冠词*the* 和*astrolabe* 联系起来，是很不自然的，也是无效的——我们在这里假设，*get* 的联想已经被正确地固定下来。在经过多次语言经历之后，机器人应该很少有机会把定冠词 *the* 与任意指示符号联系起来。

为了使这样的学习具体化，不难阐述许多不同的模型。我们只限于考虑能带来最显著特征的一个模型。

背景认知与感知假设。在明确地阐述所使用的联想和指示的学习原则之前，我首先非正式地声明，我们作出的关于这类机器人的认知能力和感知能力的假设，尽管至今还相当有限，但是，我们能与机器人一起工作。

内部语言。机器人有完全开发好的不用学的内部语言。在我们的情况下，这种语言是 Lisp 语言，这点在技术上是重要的，但在概念上却并不是基本的。当我们在这里说内部语言时，我们只是指内部表征的语言，它本身是一种更高层次的抽象语言，与机器人的具体运动和感知相关。正是在记忆中保存的内部表征语言提供了自然语言学习的直接界面。事实上，通过只用内部表征语言的机器人的行为仿真，就能进行给定自然语言的大多数机器学习。首先要学习的联想是，在强制行动记忆中的内部表征和被学习的自然语言中接近的口令之间的联想。

这种强制行动的内部表征的根本重要性，能够通过考虑动物学习的类似案例得到认可。当我们通过预期行动的引导或某种相关技巧，训练一只狗，*Get the paper*（捡起这张纸）或 *Get the ball*（捡起这个球）时，在我们称之为强制行动的记忆中的剩余记忆确实是从被证实的预期行为的丰富的感知语境中彻底地抽象出来的，而且，正是记忆中那种抽象的内部表征必须与口语刺

激相关,才能使这条狗在以后听到这个口令时执行所预期的行动。我们离知道在这种行动的记忆中狗的内部表征的更一般的结构还相差甚远。在这种有限的意义上,与机器人生活更容易得多,因为我们自己创造了机器人的内部表征形式。

对象、关系和特性。此外,我们假设,机器人从它具有的基本认知概念和感知概念开始它的自然语言学习。换言之,我们的第一语言学习实验是纯语言学习。对新概念的任何学习都被推迟到另一个阶段。例如,我们假设,机器人已经知道,所有语言中或至少在我们详细考虑的所有语言中经常提到的空间关系。这与人类的语言学习完全相反。例如,可能在不常用的起码的自然语言中,三十六个月的婴儿确实使用或完全理解左右关系。为了避免误解,我们强调,我们认为未来的一项重要任务是,让机器人也学习熟悉的时空关系。

行动。我们刚才关于对象和关系所进行的讨论也适用于行为,在英语中用像 pick up, get, place 等之类的动词来表示。当然,必须学习英语,而不是潜在的行动。

联想和文法形式。在陈述任何正式的学习原理之前,我尽可能非正式地和直观地描述可能用到的学习方案。考虑英语命令 Pick up the screw(捡起螺丝钉),其中的任何一部分机器人至今都没有学会。学习步骤可以扼要地示意如下:

(i) 通过强制或强制仿真,机器人在内存里产生了捡起螺丝钉的强制行动的一种内部表征。就学习原理的陈述而言,我们表明,这种内部表征,不是作为一种 Lisp 表达,而只是作为 Lisp 表达中的指示术语的示意性函数 I (…)。这里,我们所说的**指示术语**是指,在所提到的行动、对象、特性和关系的内部语言中的名称。于是,Pick up the screw 的内部表征是 I ($\$p$, $\$u$, $\$s$),这里 $\$p$=捡起的行动,$\$u$=方向向上,$\$s$=螺丝钉。

(ii) 通过接近联想,机器人把口令与内部表征联系起来。

$$Pick\ up\ the\ screw \sim I\ (\$p,\ \$u,\ \$s),$$

这里"～"是用来表示联想的符号。

(iii) 通过概率联想,机器人把内部符号与英语单词联系起来,一种可能性是下列不正确的结果:

$$pick \sim \$s,\ up \sim \$p,\ screw \sim \$u。$$

我们需要看出下列问题:

(a) 我们假设,从一开始,机器人就知道词的边界,像打字输入所描述的一样。这是一个假设的例子,这个例子对机器人来说是很自然的,但对于婴儿来说显然是错误的。

423

(b) 就我们的这个简单例子而言,在英语表达中,这三个内部符号与四个指示词相关联的可能方式有二十四种。我们最初赋予二十四种可能性中的每种可能性以相等的概率,但正如后面会详细说明的那样,当试验继续进行时,通过指示值的动态变化,修改了这个概率。

(iv) 建立联想之后,根据归纳原理(我们称之为范畴归纳)为每个词分配它能联想到的内部符号的范畴。在当前情况下,$pick \in O$——对象范畴,$up \in A$——行动范畴,$screw \in R$——关系范畴。于是,从这个口令也生成了一种文法形式:

$$O\ A\ the\ R$$

对英语来说,像被分配的范畴一样,这种文法形式是错误的,但记住,这只是学习的出发点。因为这种文法形式与描述其意义的内部表征 $I\ (A,\ R,\ O)$相关。

(v) 下一步发出一个新的命令,比如说,Pick up the nut。通过强制,创建了内部表征 $I\ (\$p,\ \$u,\ \$n)$ [参见上述(i)]。机器人于是首先检索其记忆,查看说出的词中的任何一个是否

与一个内部指示相关。这里,结果是 $up \sim \$ p$,而且,也发现了作为非指示词的 the 的分类。于是,pick、the 和 nut 的概率联想的可能性有六种。注意,这里不再出现前面 pick 与 $\$ s$ 的不正确的联想,这意味着,在这个学习阶段,联想发生了变化。因此,让我们假设,新的联想是

$$pick \sim \$ u,\ nut \sim \$ n。$$

作为一种新的文法形式我们也有

$$R\ A\ the\ O$$

这种文法形式尽管不正确,但现在它只源于混淆了 pick 和 up 的联想。为了纠正这些联想,我们必须拆开 pick 和 up 的固定搭配,这正是我们要做的。无论如何,我们立刻形成了内部表征的联想:

$$R\ A\ the\ O \sim I\ (A,\ R,\ O)。$$

(vi) 向机器人发出一个口令,只要能成功地完成解释的下列步骤,学习就停止:

(a) 对于每个词来说,在记忆中发现了一个内部指示或非指示分类的一种联想。

(b) 在记忆中发现了每个词的范畴。

(c) 根据记忆中的一个相关的内在表征,发现了由(b)产生的文法形式。

(d) 基于这个内部表征,正确地执行了命令。

我现在转向这些思想的详细进展,从内部语言开始,然后是学习公理。 424

内部语言。我用 Lisp 表示这里所报告的这种研究的内部语言。内部语言先于学习存储在记忆中,在学习期间不会发生任何变化。

在表 1 中，通过文法指定了表达内部语言的集合，其中，词汇范畴 A_1，A_2，A_3，A_5(=行动)，REL (=关系)，PROP (=特性)，OBJ (=对象特性)和短语范畴 A (=行动)，S(=对象集)，O (=对象)，G (=区域)，以及 DIR (=方向)。表 2 给出了我们的内部语言的词汇。我们把词汇范畴的元素看成是**内部符号**。运算，比如 fa_1 和 fa_2，读作**形式行动**，在一种给定的机器人环境中，所有这些操作都有一个直接的程序解释。

表 1　内部语言的文法

Ⅰ	*A*	→ $(fa_1\ r_1\ O)$
Ⅱ	*A*	→ $(fa_2\ r_2\ G)$
Ⅲ	*A*	→ $(fa_3\ r_3\ O G)$
Ⅳ	*A*	→ $(fa_5\ r_5\ DIR\ O)$
Ⅴ	*A*	→ $(fa_5\ r_5\ O)$
Ⅵ	*G*	→ $(fr\ REL\ O)$
Ⅶ	*DIR*	→ $(fd\ REL)$
Ⅷ	*O*	→ $(io\ S)$
Ⅸ	*O*	→ $(so\ S)$
Ⅹ	*S*	→ $(fo\ PROP\ S)$
Ⅺ	*S*	→ $(fo\ OBJ\ *)$

表 2　内部语言的词汇

Categories							Semantic Operations
OBJ	PROP	REL	A_1	A_2	A_3	A_5	
$screw	$large	$up	$get	$go	$put	$pick	fa_1 (form-action)
$nut	$medium	$on			$place		fa_2
$washer	$small	$into					fa_3
$hole	$square	$above					fa_5
$plate	$hexagonal	$to					*fr* (form-region)
$sleeve	$round	$behind					*fdir* (form-direction)
	$black						*io* (identify-object)
	$red						*so* (select-object)
	$gray						*fo* (form-object)

表 2 中所用的英语词反映了一个英语词汇，但内部语言的句法是 Lisp，不是英语。我们的范畴与常用的语言范畴十分匹配：A 对应于一个（祈使）句范畴，A_1 对应于及物动词范畴，REL 对应于介词范畴，PROP 对应于形容词范畴，OBJ 对应于普通名词范畴，DIR 对应于副词范畴，G 对应于介词短语范畴，O 对应于（确定的）名词短语范畴，以及 S 对应于一个名词词组的范畴。然而，我们选择通过这些范畴的语言标记不指称这些范畴，因为我们把它们看成是语义范畴。 425

这种内部语言文法将推导出对应于英语命令 Get a screw 的行动的内部表征的 Lisp 结构，这里，星号（∗）指在一个特定视觉环境中出现的对象集：

$$(fa1\ \$\, get\ (so\ (fo\ \$\, screw\ *)))。\qquad (1)$$

设 $\gamma = (fo\ \$\, screw\ *)$。于是，$\gamma$ 本身是(1)中只包含内部符号 $\$\, screw$ 的**最小的** Lisp 表达，$(so\ (fo\ \$\, screw\ *))$ 是(1)中只包含内部符号 $\$\, screw$ 的**最大的** Lisp 表达。我们后面用到了这个区分。

一般学习公理。我们现在转向我们的学习公理，这些公理自然分成两组，一组代表运用工作记忆的计算，另一组代表在长期记忆的状态中的变化。显然，我们用了对记忆类型的一种区分，这种区分在人类记忆的心理学研究中是标准的，但是，我们的机器学习过程的细节不必忠实于人类学习语言的过程，我们也不要求它们应该如此。另一方面，我们的联想、概括、分类和规则生成的基本过程，在人类学习中，几乎肯定有类似情况，当前，一些基本过程比另一些基本过程更好地得到理解。在本节阐述的一般公理中，我们假设内部表征的特殊语言相当少，尽管举例说明这些公理的例子用了刚才描述的内部语言。

符号。至于公理中用到的符号，一般情况下，只要是自然语言，我们就用拉丁字母表示句子或句子的组成部分，而且，我们用希腊字母指句子或句子的组成部分的内部表征。现在转向特殊符号，字母 a, b, …… 指一个句子中的词，希腊字母 α, β, ……指内部符号。符号 s 指一个完整的句子，相应地，σ 指一个完整的内部表征。文法形式——句子形式或术语形式——用 g 表示，或者也用 $g\ (X)$ 表示，以表明一种形式的范畴自变量；相应地，文法形式的内部表征用 γ 或 $\gamma\,(X)$ 表示。在语义范畴或范畴变量 X、X'、Y 等情况下，我们违反了我们的希腊-拉丁字母的约定。我们在文法形式及其内部表征中用了同样的范畴符号。

为了确信把正确的语义含义从一个自然语言的句子带到它的内部表征，或反之亦然，我们根据需要对一个给定句子中多次出现的相同范畴和在其内部表征中的相应出现作出索引。后面举了这种索引的一个例子。

426

运用工作记忆的计算[①]

W1. 概率联想。在任意试验中，设由 s 联想到 σ，设 a 在与 σ 的任意内部符号无关的 s 的词的集合中。再设 α 在当前与 s 的任意词无关的内部符号的集合中。于是，抽样对 $(a,\ \alpha)$，可能运用了当前的指示值和关联，即 $a \sim \alpha$。

在 Get the screw 的情况下的概率抽样可能导致不正确的联想 *get*～＄*screw*、*the*～＄*get*，而且，没有关于 *screw* 的联想，因为在内部表征中只有两个符号是关联的。

---

① 学习程序有处理所出现的命令的一种短期记忆。这种记忆保留了它在一次试验期间的内容。第一组学习公理描述在试验期间发生在工作记忆中的联想计算。

$W2$. 形式概括。[①] **如果 $g\ (g_i')\sim\gamma\ (\gamma_i')$，$g_i'\sim\gamma_i'$和 γ' 可从 X 中导出，那么 $g\ (X_i)\sim\gamma\ (X_i)$，这里 i 是发生事件的索引。**

从公理 $W1$ 之后给出的联想来看，我们会推导出不正确的概括

$$OBJ\ A_1\ screw\sim(fa_1\ A_1(io\ (fo\ OBJ\ \ *\ )))。\qquad(2)$$

正确的概括是

$$A_1\ the\ OBJ\sim(fa_1\ A_1(io\ (fo\ OBJ\ \ *\ )))。\qquad(3)$$

$W3$. 文法规则生成。如果 $g\sim\gamma$ 和 γ 可从 X 中导出，那么 $X\rightarrow g$。

与 $W3$ 相对应，我们现在得到不正确的规则

$$A\rightarrow OBJ\ A_1\ screw。\qquad(4)$$

正确的规则是

$$A\rightarrow A_1\ the\ OBJ\ 。\qquad(5)$$

$W4$. 形式联想。如果 $g\ (g')\sim\gamma\ (\gamma')$ 和 g' 以及 γ' 有相应的被索引的范畴，那么 $g'\sim\gamma'$。

我们从(2)中得到不正确的形式联想

$$OBJ\sim\ (io\ (fo\ OBJ\ \ *\ ))。\qquad(6)$$

正确的联想——从更多的试验中学到的——是从(3)中导出的

---

① 一个明确的概括原理产生了文法形式及其联想的语义形式。所用的方法把语境无关文法的那些方法与形式语言的模型理论语义学结合起来。概括概念在心理学的学习理论中得到了广泛的运用。这里所用的这种概括限于从具体表达中产生的文法形式和文法规则。例如，短语 *the nut* 概括为文法形式 *the O*，这里 O 是对象范畴。

$$the\ OBJ \sim (io\ (fo\ OBJ\ \ *))。\qquad (7)$$

*W*5. 形式规范。如果 $g\ (X_i) \sim \gamma\ (X_i)$，$g' \sim \gamma'$ 和 γ 可从 X 中导出，那么 $g\ (g'_i) \sim \gamma\ (\gamma'_i)$。

正如能容易看到的那样，关于形式规范的公理 *W*5 实质上是关于形式概括的公理 *W*2 的逆命题。

*W*6. 内容删除。在每次试验结束时删除工作记忆的内容。

427 我们现在转向支配长期记忆的公理。

在长期记忆状态中的变化①

*L*1. 指示值计算。如果在试验 n 结束时，在当前语言刺激中由一个词 a 联想到某个内在符号 α，那么 a 的指示值 $d(a)$ 增大，如果 a 没有这样的联想，$d(a)$ 就减小。此外，如果一个词 a 在试验中没有出现，那么 $d(a)$ 保持不变，除非在试验中破坏 a 与内在符号 α 的关联，在这种情况下，$d(a)$ 减小。

因为这个公理在概念上不太熟悉，所以，我们后面举一个更详细的例子。

*L*2. 形式因式分解。如果 $g \sim \gamma$ 和 g' 是已经在长期记忆中的 g 的子字符串，并且，g' 和 γ' 可从 X 中导出，那么 g 和 γ 可被还原为 $g\ (X)$ 和 $\gamma\ (X)$。同样，$g\ (X) \sim \gamma\ (X)$ 被存储在长期记忆中，正如公理 *W*4 产生的相应的文法规则那样。

我们举一个例子来说明这个公理，这个例子其实很简单。它看起来复杂，因为前提占了三行，并且，我们有两种结论：一种联想和对应的文法规则。设

---

① 这种记忆也会随着试验的变化而变化，但是，当联想和文法形式对所考虑的应用来说是正确的时，它存储了保持不变的联想和文法形式。在本节的学习公理部分描述了从一次试验到另一次试验长期记忆状态发生变化的方式。

$$g \sim \gamma: A_1\ the\ OBJ \sim (fa_1\ A_1(io\ (fo\ OBJ\ \ *\ )))$$
$$g' \sim \gamma': the\ OBJ \qquad \sim (io\ (fo\ OBJ\ \ *\ ))$$
$$X: O \qquad \rightarrow the\ OBJ$$

那么

$$A_1O \sim (fa_1A_1O)$$
$$A \rightarrow A_1O。$$

L3. 形式过滤。我们随时可以从长期记忆中消除联想和文法规则,如果它们能够被生成的话。

在前面的例子中,从作为公理 $W3$ 的一个例子了解到,现在能从长期记忆中消除 $g \sim \gamma$,也能消除 $A \rightarrow A_1\ the\ OBJ$ 。

L4. 全等计算。[①] **如果 w 是 g 的一个子字符串,w' 是 g' 的一个子字符串,并且它们是这样的:**

(a) $g \sim \gamma$ 和 $g' \sim \gamma$;

(b) 只有在发生 w' 取代 w 时,g' 不同于 g;

(c) w 和 w' 不包括高指示值的词。

那么,$w' \approx w$,而且全等存储在长期记忆中。

运用公理 $L4$,一个给定自然语言的文法规则数,通过运用意义全等,进一步得到减少(Suppes, 1973b,1991)。考虑下列文法形式的联想: 428

---

① 通过运用限于非指示词的一个语义全等概念,我们简化了这些文法,同时,允许进行跨语言的直接比较。这种全等的直观思想是简单的。当一种自然语言的两个字符串在内部语言中有完全一致的表征时,它们是全等的。

我们能够运用不同强弱的全等概念,得到意义的不同相似度。在寻找一个同义词概念时,不会有这种想法,就像我们不会在一个全等概念中捕获到现代几何一样。例如,我们在仿射几何学中具有的全等意义比欧几里得几何学中的全等意义更弱,但也很容易得到一个比欧几里得几何学中的全等概念更强的通常的全等概念,即要求方向相同的全等。

$$die\ Schraube \sim (io\ (fo\ \$\ screw\ *\ )) \tag{8}$$

$$der\ Schraube \sim (io\ (fo\ \$\ screw\ *\ )) \tag{9}$$

联想(8)和(9)只有冠词不同。在(8)中的冠词属于主格和宾格的情况,在(9)中的冠词属于所有格和与格的情况。这里重要的是,各自的内部表征没有差别。我们因此称(8)与(9)全等,并且,把不同元素收集为一个全等类[*DA*]={*die*, *der*},这里 *DA*=定冠词。这允许我们把两种文法形式(8)与(9)简化为一种:

$$[DA]\ Schraube \sim (io\ (fo\ \$\ screw\ *\ ))。 \tag{10}$$

注意,这种根据全等的简化在下列方式中是有风险的。我们可能失去所学语言的信息。例如,摧毁(8)和(9)之间的差别显示出的不同性别,将使我们不能辨别下列句子:

Steck die Schraube in das Loch. (11)

Steck die Schraube in die Loch. (12)

据此(11)是文法的,(12)不是。只要我们的焦点是关注理解性文法,像(12)那样的一个命令就不可能产生,但为了达到产生式的目的,就不应该以当前的形式使用全等。

*L*5. **记忆痕迹的形成。第一次形成一种形式概括、文法规则或全等时,作为这种概括、文法规则或全等的基础的词联想,与它一起被存储在长期记忆中。**

运用我们在 *W*3 之后最初的例子,不正确的联想会被存储于长期记忆中,但随着更多的学习,以后会被删除[*L*6(a)]。

*L*6. **联想的删除。**①

---

① 就更广泛的应用而言,这个公理太强。在自然语言的许多熟悉的运用中,许多联想是有意义的。这里,我们彻底地把这些联想简化为严格的语义联想。这种高度限制的简化,与本章前面几节一般使用的多种计算联想相反。对限于简单化的惟一辩护是,在这个早期发展阶段,使机器学习中的工作更容易。

(a) 当一个句子中的一个词被赋予一种新的联想时，这个词前面的任何联想都会被从长期存储器中删除。

(b) 如果在一个试验开始 $a \sim \alpha$，a 出现在这次试验给出的表达 s 中，但 α 没有出现在 s 的内部表征 σ 中，那么联想 $a \sim \alpha$ 就被从长期记忆中删除了。

(c) 如果没有从一个句子 s 的出现生成任何一种内部表征，那么 σ 是作为正确的内部表征给出的，并且，在 s 中有几个词与 σ 的内部符号 α 相关，使得这些词的出现次数大于 α 在 σ 中的出现次数，那么这些联想被删除了。 429

*L*7. 形式联想或文法规则的删除。如果 $a \sim \alpha$ 被删除，那么任何一种形式概括、文法规则或全等（对此而言，$a \sim \alpha$ 是一种记忆痕迹）也被从长期存储器中删除。

对公理的评论。在十三个公理中，只有三个公理需要在这里进行更详尽的报道研究。这三个是，(*W*1)概率联想、(*W*4)形式联想和(*L*1)指示值计算，下一小节对此作了更具体的技术性表述。特别对公理 *W*2 进行了更详细的表述。也为这些更具体的公理举出了一些例子。

某些公理和初始条件的详细说明。

概率联想。(公理 *W*1)在任意试验 n 中，设 s 与 σ 相关，设 A 是与 σ 的任意内部符号无关的 s 的词的集合，设 $d_n(a)$ 是 A 中每个这样的 a 的当前指示值，再设 $\mathbb{A}$ 是当前与 s 的任意词无关的内部符号的集合。那么

1. 一个元素 α，如果没有从 $\mathbb{A}$ 中置换，则被均匀抽样。

2. 同时，一个元素 a，如果没有从 A 中置换，则以下列抽样概率被抽样，

$$p(a) = \frac{d_n(a)}{\sum_A d_n(a)}。$$

3. 被抽样的对是相关的，即 $a \sim \alpha$。

4. 一直进行抽样直到集合 A 或内部符号的集合 $\mathbb{A}$ 是空的时为止。

由于这个程序的概率本质(公理 W1)，存在着几种可能的结果。例如，考虑 Get the screw，它有内部表征(fa_1 \$*get* (*io* (*fo* \$*screw* *)))。这个抽样过程可能生成六个不同可能的对中的任何一个对，比如，像 *get*~ \$*screw* 和 *screw*~ \$*get*。既然在口令中有三个词出现，原则上，有六种可能方式把这个命令的三个词与两个内部表示符号联系起来。

形式联想。(公理 W4′)设 $g \sim \gamma$ 是任意试验中一种联想计算的任意一步。

1. 如果 X 在 g 中出现，(*fo* X *)在 γ 中出现，那么 $X \sim$ (*fo* X *)。

2. 如果(i) wX 是 g(其中，$g \sim \gamma$)的一个子字符串，使得 $w = a$，这是一个有低指示值的词，或者，如果 X 的前面是一个变量，或者是 g 的第一个符号，那么 $w = \varepsilon$(空符号)，而且，(ii) $\gamma'(x)$ 是含 X 发生和没有其他范畴发生的的最大 Lisp 形式，那么

$$wX \sim \gamma'(X)。$$

430 3. 如果(i) $X_1 w_1 \cdots w_{m-1} X_m$ 是 g 的一个子字符串，这里，X_i，$i=1, \cdots, m$ 未必都是不同的范畴名称，w_i 是可能为空的子字符串，或者是与特定试验中的内部符号无关的词，而且，(ii) $\gamma'(X_{\pi(1)}, \cdots, X_{\pi(m)})$ 是含有 $X_{\pi(1)}, \cdots, X_{\pi(m)}$ 的 γ 的最小 Lisp 形式，那么

$$X_1 w_1 \cdots w_{m-1} X_m \sim T(X_{\pi(1)}, \cdots, X_{\pi(m)}),$$

这里 π 是数字 $1, \cdots, m$ 的一种排列。

为了表明公理 W4′如何起作用，假设我们已经获得下列文

法形式的联想：

$$A_1\ the\ PROP\ OBJ \sim (fa_1\ A_1(io\ (fo\ PROP\ (fo\ OBJ\ \ *))))\tag{13}$$

这可能作为一种概括，比如，根据有正确联想词的命令 Get the red screw 来获得。为了分析这个例子，我们参考表 1。

根据公理 $W4'.1$，我们可以推出

$$OBJ \sim (fo\ OBJ\ \ *)。\tag{14}$$

根据公理 $W4'.2$，我们推出

$$PROP\ OBJ \sim (fo\ PROP\ (fo\ OBJ\ \ *))。\tag{15}$$

根据公理 $W4'.3$，我们推出

$$the\ PROP\ OBJ \sim (io\ (fo\ PROP\ (fo\ OBJ\ \ *)))。\tag{16}$$

运用文法规则生成（公理 $W3$）和内部语言的文法（表 1），我们根据(14)和表 1 的规则 XI 推出

$$S \rightarrow OBJ。\tag{17}$$

根据(15)、表 1 的规则 X 和形式概括（公理 $W2$）推出

$$PROP\ S \sim (fo\ PROP\ S),\tag{18}$$

最后根据文法规则生成（公理 $W3$）推出

$$S \rightarrow PROP\ S\tag{19}$$

作为英语文法的一个规则。我们运用公理 $W2$，也根据(16)、(15)和内部文法推出

$$the\ S \sim (io\ S)\tag{20}$$

然后，再一次根据文法规则生成推出

$$O \rightarrow the\ S\tag{21}$$

作为我们的英语文法的一个规则。

在引入这些公理之前，为了保持意义，我答应举一个范畴索引的例子。对于这里所限定的语料库来说，能够避免这样的索引，但就更一般的目的而言，这是需要的。下面是从语料库中举出的一个例子，表明这种索引是如何进行的。考虑下列句子，

Put the nut on the screw.

431 根据索引，正确的文法形式和相关的内部表征看起来像是这样的，

$$A_3\ the\ OBJ1\ REL\ the\ OBJ2\sim$$

$$(fa_3A_3(io\ (fo\ OBJ1*))\ (fr\ REL\ (io\ (fo\ OBJ2*)))),$$

这里，写在后面的数字用来索引 OBJ。

指示值计算。（公理 $L1'$）如果在试验 n 结束时，在所呈现的言语刺激中的一个词 a 与 s 的内部表征 σ 的某个内部符号 α 相关，那么①

$$d_{n+1}(a)=(1-\theta)d_n(a)+\theta$$

如果 a 与这个内部表征的某个所指的内部符号 α 无关，

$$d_{n+1}(a)=(1-\theta)d_n(a)$$

而且，如果一个词 a 在试验 n 中没有出现，那么

$$d_{n+1}(a)=d_n(a)$$

除非在试验 n 中，破坏了 a 与内部符号 α 的关联，在这种情况下

$$d_{n+1}(a)=(1-\theta)d_n(a)$$

为了表明如何计算指示值的计算（公理 $L1$），让我们假设，给定

---

① 这里所用的学习模型只是上一节(4)开始时引入的线性学习模型。

的联想是 $get \sim \$ screw$，$the \sim \$ get$。让我们进一步假设，在这次试验结束时，

$$d(get)=0.900$$

$$d(screw)=0.950$$

$$d(the)=0.700。$$

在下一次试验中，口令是

Get the nut.

结果，我们以

$$get \sim \$ get,\ nut \sim \$ nut$$

和删除 the 的联想[公理 L6 (a)]，结束了这个试验，运用 $\theta=0.03$，正如我们通常所做的那样，我们现在有

$$d(get)=0.903$$

$$d(the)=0.679。$$

之后，让我们说，the 再多出现三次，也没有形成任何联想，指示值将进一步减小到 0.620。如果给出命令 Get the sleeve，而且，sleeve 在前面没有出现过，以及 $get \sim \$ get$，那么，我们可以看到，sleeve 与 \$ *sleeve* 相关的概率如何比与 the 相关的概率更高。在所给出这些假设条件下，the 和 sleeve 的抽样概率将是

$$p(the)=\frac{d(the)}{d(the)+d(sleeve)}=\frac{0.620}{0.620+1}=0.383$$

以及 432

$$p(sleeve)=\frac{d(the)}{d(the)+d(sleeve)}=\frac{1}{1.620}=0.617。$$

在初始学习之后，指示值的动态计算一直进行，即使不发生错

误。结果,像英语中的 a 与 the 和汉语中的 ba 那样的高频词汇,它们的指示值很快趋于零,正如从稍后给出的图 1 和 2 中的学习曲线可以看出的那样。(从一种形式的观点来看,如果一个词的渐进指示值为零,或者更实际地说,低于某个阈限,把它定义为**非指示的**,是有用的。)

这个有两个参数 $d_1(a)$和 θ 的特殊的线性学习模型,可能很容易被更精致的替代模型所取代。

初始条件。在试验 1 开始时,联想关系~、全等关系≈和文法规则的集合都是空的。此外,初始指示值 $d_1(a)$对于所有的词 a 都是一样的。

语料库。为了检验我们的系统,我们把它应用于十个语料库,每个语料库都有一种不同的语言。这些语言是:英语、荷兰语、德语、法语、西班牙语、加泰罗尼亚语、俄语、汉语、韩语、日语。① 我们的语料库大小从四百个句子到四百四十个句子不等。在十种语言中的语料库涵盖一个几乎完全相同的内部结构的集合。由于下列原因,它们不可能完全相同:一个内部语言的词,在一种语言(比如 $\mathcal{L}$)中用一个词来翻译,在另一种语言 $\mathcal{L}$ 中可能需要两个或多个词(依赖于语境)来翻译。结果,从四百个句子的语料库中,$\mathcal{L}$ 是不可能学会的。为了达到完全学会所有词的目的,我们因而要么增加句子,比如,像在西班牙语中那样,要么像在日语中那样删除句子。

在不同的语料库中最重要的变化的必要条件是,一种给定语言中的两个词,旨在对应于两个内部符号,如果能了解正确含

---

① 这个排序反映了对语言接近程度的定性判断,正如在下面的表 5 中可以明确地看出的那样,它表明了每种语言的理解性文法规则。例如,由于把命令式的行为动词放在最后,所以,把汉语、韩语和日语一起分为一组。

义,那么它们未必总是同时出现。例如,如果 *nimm* 和 *Schraube* 只在一个命令 *Nimm die Schraube!* 中出现,就不能保证能学会所预期的联想 *Schraube*～＄*screw* 和 *nimm*～＄*get*,不管进行多少次学习试验。

我们也注意到,在日语的情形中,我们删除了翻译作为一个内部符号＄*above* 的所有句子,因为在日语中,*above* 和 *on* 用同一个词 *ue* 表示。因此,这意味着我们在日语的情形中要从语料库中删除十个句子。

尽管我们的译者作出了周到的讲解,我们还是不能确定,这些翻译是否在任何情况下都听起来很自然,并确实能用在机器人的学习环境中。然而,核实这一点大大超越了标准翻译方法能做的事情,并要求在一种工作环境中对语言的运用进行现场研究。在词汇空缺的情况下,我们运用很简单的方法:例如,法语没有一个惯用词表示英语形容词 *medium* 的非空间含义,因此,我们用冗长迂回的短语 *de taille moyenne*。在某些情况下,我们避免技术性语言,在这里,一个词由两个语素组成并表示两个内部指示符号。如期所料,这经常发生在德语中。例如,*Rundschraube* 表达了 *round screw* 的思想,也就是,特性 *round* 和对象 *screw*。

433

在加泰罗尼亚语的情形中,我们事先把一些词的指示值设定为不同于其初始值 1。这样做的理由是,在我们的意义上,有如此多的非指示词,并且这些非指示词一律随着某些行为动词同时出现,所以,在我们有限的语料库中,只能以小于 1 的概率学习直观上正确的指示值。因此,我们把词 *d* 和 *de* 的初始指示值设定为 0.05。

为了在一个语料库中获得成功的学习绩效,通常有用的是,在这个语料库中做一些调整。当然,在对一个理论的最后检验

中,这样的改变当然是不希望有的。下面描述了我们作出的改变。

由于 pick up 这个动词的特殊本性和在英语中的关系,它的引入带来了许多问题。在许多语言中,用一个不必附加介词的行为动词来表达 picking something up 的概念。正如在不同语言中所表达的那样,由这种特殊行为产生的这些问题,导致了一个重要的问题。在一种特定的自然语言的语料库中,个别词有时不止指一个内部符号。我们的解决方案是把这些词分成两部分,这个方案是人为的,但我们不得不采用真正确切的人为解决方案:

法语:*ramasse* 分成 *ra* 和 *masse*

西班牙语:*recoge* 分成 *rec* 和 *oge*

加泰罗尼亚语:*recull* 分成 *re* 和 *cull*

俄语:*podnimi* 分成 *pod* 和 *nimi*

韩语:*cip&ela* 分成 *cip* 和 *&ela*

日语:*toriagero* 分成 *tori* 和 *agera*

正如在鲍尔曼(1996)以及乔伊(Choi)和鲍尔曼(1991)的文章中所观察到的那样,跨语言词汇的语义范畴是不同构的,这并不令人感到惊奇。在一种语言中被认为是相同的行动和不同的关系,可能在另一种语言中被认为是不同的行动和相同的关系。

经验结果。在这一小节总结我们的结果时,我们首先提供一些学习曲线,随后,在我们的意义上,介绍非指示词的全等分类。我们然后运用这些抽象分类来简化和提供在语义上从十种语言生成的理解的基本文法规则的汇总表。

词汇。在表 3 中,我们为每一种自然语言规定被学习的词的个数。我们把词典用法中是相同的词的不同变形看成是不同

的词汇。如期所料，俄语，正是由于它丰富的变形模式，具有不止一种联想的内部符号数最多。

表3　十种可比较的语料库的理解词表

434

	词	非指示词	多于1种联想的符号
英语	28	2	0
荷兰语	38	3	2
德语	49	10	9
法语	42	7	5
西班牙语	40	7	6
加泰罗尼亚语	42	8	8
俄语	75	0	16
汉语	33	8	0
韩语	31	4	1
日语	31	4	1

学习。一个自然的问题是，为了促进学习，语料库中的这些句子的出现顺序是否是预先固定的。回答是，这个顺序不是固定的。就每一种语言而言，没有从大约四百个句子的语料库中进行替换，而是随机地选择所出现的句子。

与大多数神经网络的学习速率相比，这种学习很快。在十种语言的每一种语言中，经过整个语料库的一次循环，足以产生一种直观上正确的理解性文法。相反，在有点悖论的意义上，正如我们在苏佩斯、良和伯特纳的文章中所证明的那样，理论上基于句子的随机抽样的这种标准的平均学习曲线，即使对于只有四百个句子的一个语料库来说，在计算上也是不可行的。在同一篇较早的文章中，基于非常大的运行——事实上，由高达一万个抽样路径来计算平均学习曲线——我们推测，我们所研究的这种语料库的平均学习速率是多项式界的。就我们的理论学习

框架而言,所研究的十种语言的学习,相当明确地支持了这种推测。

在图1—3中,我们显示出英语、汉语和德语词的指示值的平均学习曲线。这种情况下的平均是关于一个给定语料库中的指示词或非指示词的总数的平均。在这三种语言中的非指示词的个数分别是2、8和10,也如表3所示。如期所料,非指示词的指示值的学习速率,对于三种语言来说,与非指示词的个数成反比,即语料库中出现的纯频率的一个论点。正如通过比较汉语和德语的平均曲线所能看出的那样,即使非指示词的个数非常接近,这也不是全部情况。

在图4中,我们显示了英语中六十个物理应用题相应的平均曲线,也是只含有两个非指示词 the 和 a (suppes and liang, 1996)。这种内部语言完全不同于机器人的情况,但除了说它在本质上完全是物理量的方程式语言之外,我们这里不打算对它作出描述。然而,我们强调,对于机器人的命令和物理应用题来

435

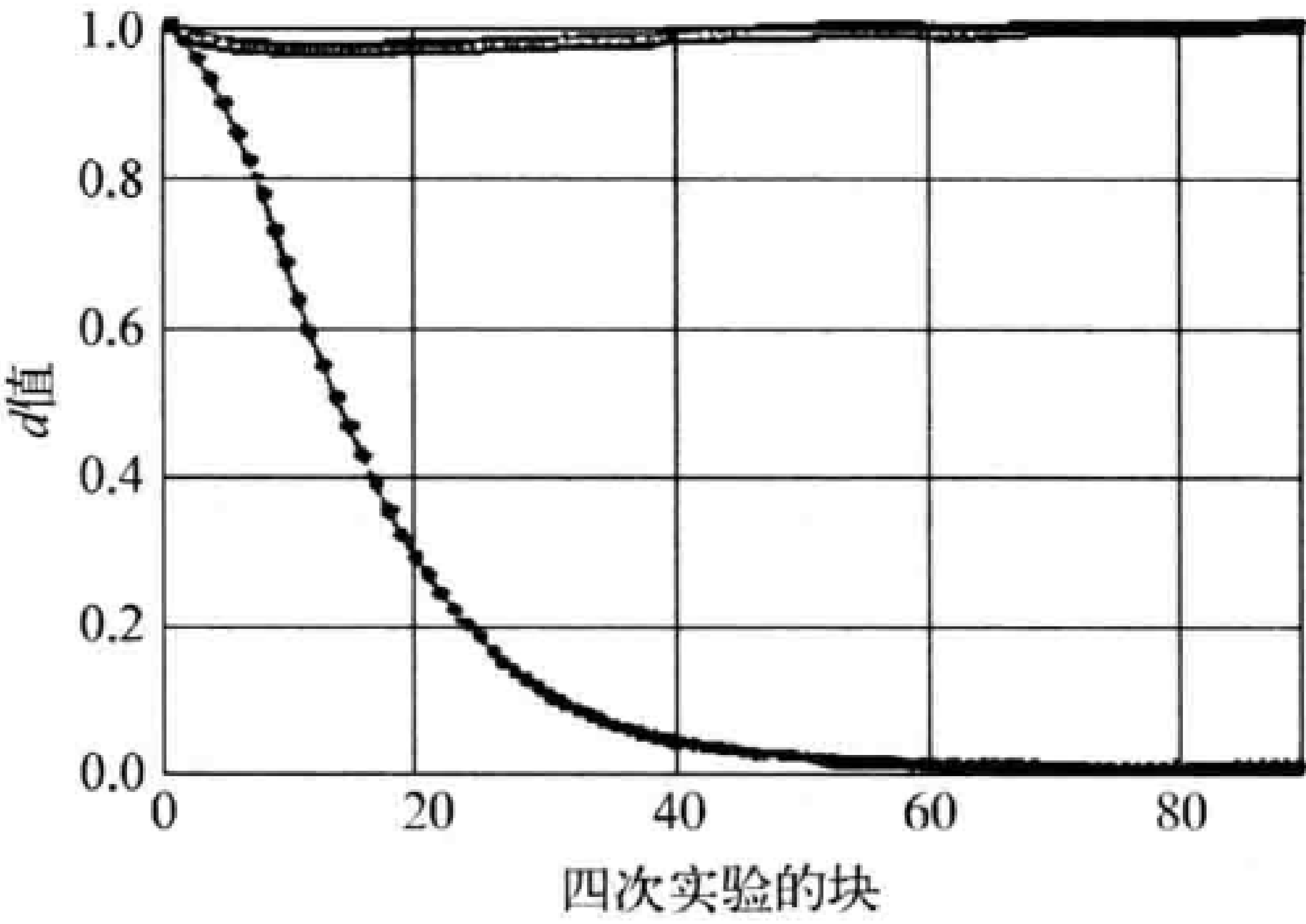


图1 英语的平均指示学习曲线。上面的曲线是表示指示词,渐进地趋于1,下面的曲线是表示非指示词,渐进地趋于0

说，学习公理是一样的。

全等类。一般而言，为了比较不同语言所运用的诸规则，我们尽可能地跨语言合并。全等类跨语言的最重要的扩展是引入空词 ϵ，以便当在一个特殊的地方没有指示词出现时，一种给定语言仍然被包括在运用那个规则的语系中。这有一种结果，即只通过这个规则中语义范畴出现的顺序，而不是通过非指示词的出现，区分这些文法规则。换言之，以完全相同的顺序出现的完全相同的语义范畴具有的两个规则，与非指示词的出现无关，被看成与下面表 5 中的规则相同。为了对不同语言生成的理解性文法作出真正的语义比较，这种全等简化是令人向往的。表 4 中出现的 ϵ 表明，这种语言运用的文法规则与另一种语言一样，但没有非指示词出现。

上面所说的一种结果恰好是，含有全等符号的一个特殊的文法规则，在语料库本身没有例示的一种文法形式的给定语言中，可能允许有一种例示。为了理解的目的，与产生式相反，这不会导致困难。

436

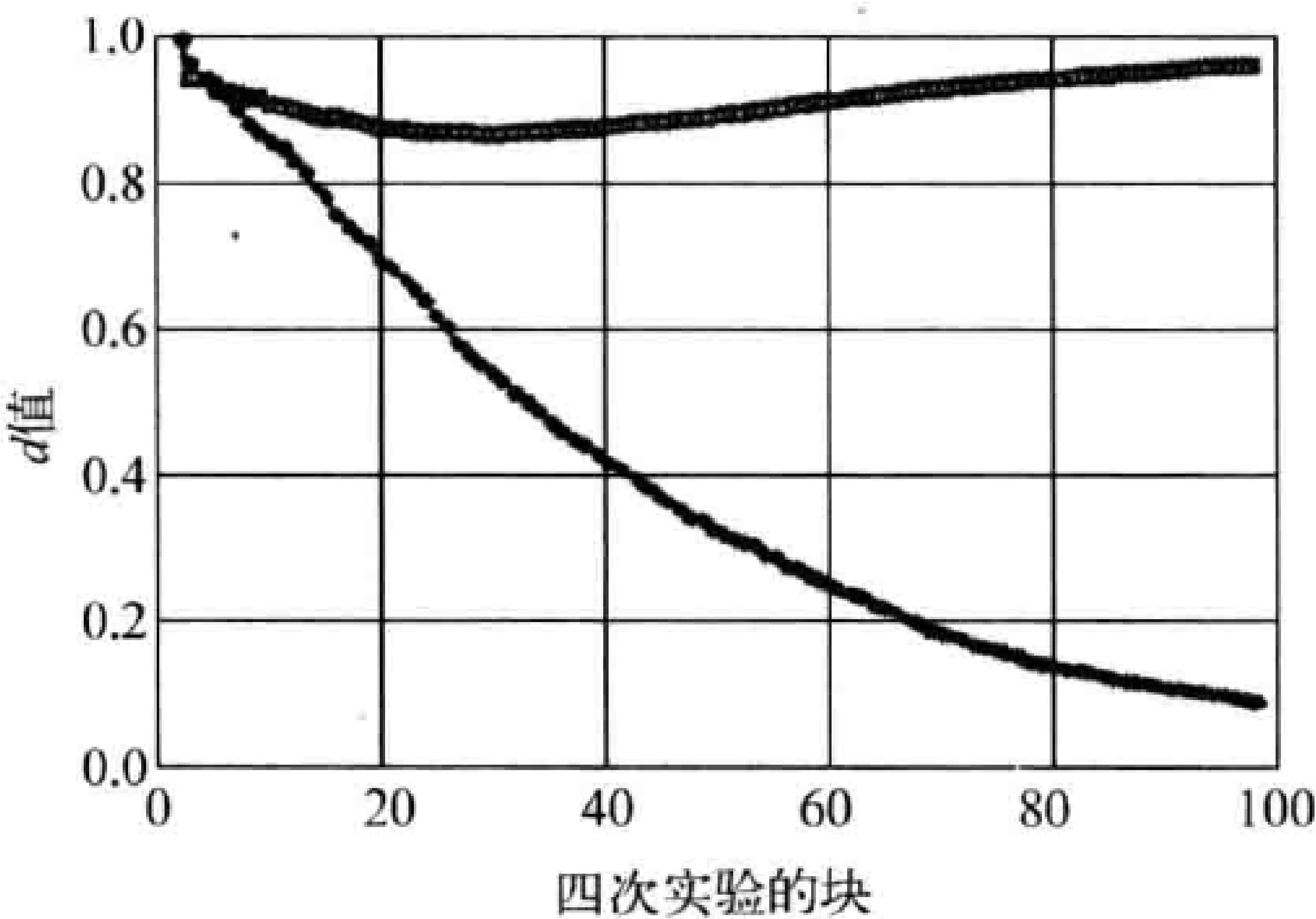


图 2　汉语的平均指示学习曲线

关于表 4 有许多描述性的评论。全等类 $\gamma1$ 是由各种形式的定冠词组成的,包括在俄语和日语情况下的空词 ϵ ,因为这两种语言没有在标准意义上运用定冠词。同样的评论适合于不定冠词的类 $\gamma2$。在跨越语言引入类 $\gamma1$ 和 $\gamma2$ 时,我们诚然引入一种粗粒的但有用的非指示词群,其功能在不同的语言中相当类似。对意义和用法的全等的更进一步的提炼可能导致把 $\gamma1$ 和 $\gamma2$ 划分成更小的类。

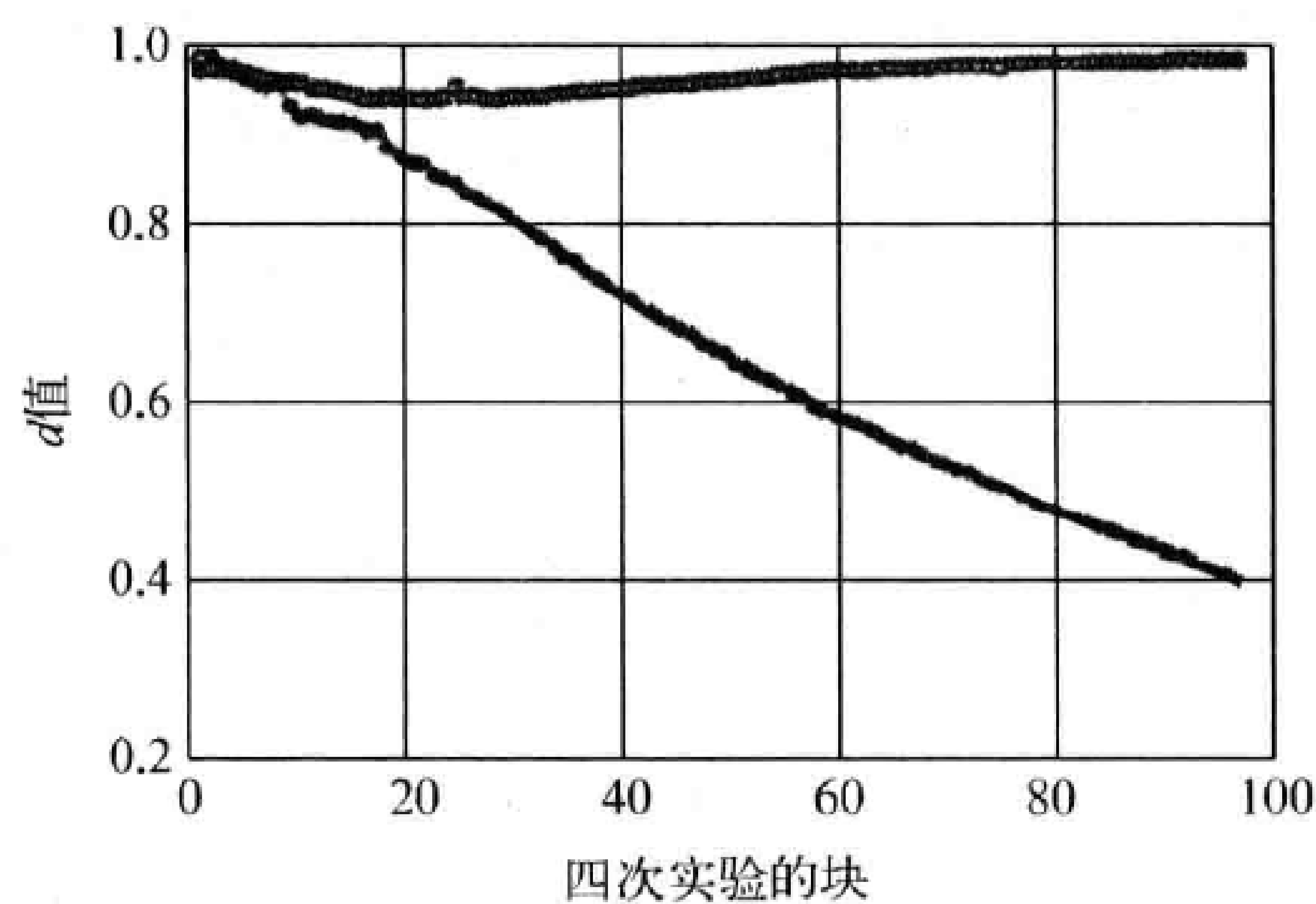


图 3　德语的平均指示学习曲线

437

表 4　全 等 类

	E	D	G	F	S	C	R	Ch	K	J
$\gamma1$	*the*	*het* *de*	*die* *das* *der* *dem* *den*	*le* *l* *la*	*el* *l* *la*	*el* *l* *la*	ϵ	*nage* *zhege*	*ku* ϵ	ϵ ϵ
$\gamma2$	*a*	*een*	*ein* *eine* *einem* *einen* *einer*	*un* *une*	*un* *una*	ϵ	ϵ	*yige*	*han* ϵ	—

（续表）

	E	D	G	F	S	C	R	Ch	K	J
γ3	—	—	—	*de taille* ϵ	ϵ	ϵ	—	—	—	—
γ4	ϵ	ϵ	ϵ	ϵ	—	—	ϵ	ϵ	ϵ	*no* ϵ
γ5	—	—	—	—	—	—	—	ϵ	ϵ	*no tokoro* *no*
γ6	ϵ	ϵ	ϵ	*de* ϵ	*de* ϵ	*d* *de* ϵ	ϵ	ϵ	—	—
γ7	—	—	—	—	—	—	—	—	—	*o*
γ8	—	—	—	—	—	—	—	*ba*	ϵ	ϵ
γ9	—	—	—	—	—	—	—	—	*ul* *lul*	*o*
γ10	—	—	—	—	—	—	—	—	ϵ	*ni*
γ11	—	—	—	—	—	—	—	*zai* ϵ	—	—
γ12	—	—	—	—	—	—	—	*chao* ϵ	ϵ	ϵ
γ13	—	—	—	—	—	—	—	*na4* *na4li* ϵ	ϵ	*ni*
γ14	ϵ	ϵ	ϵ	ϵ	*junto* ϵ	*al* ϵ	ϵ	—	—	—

438

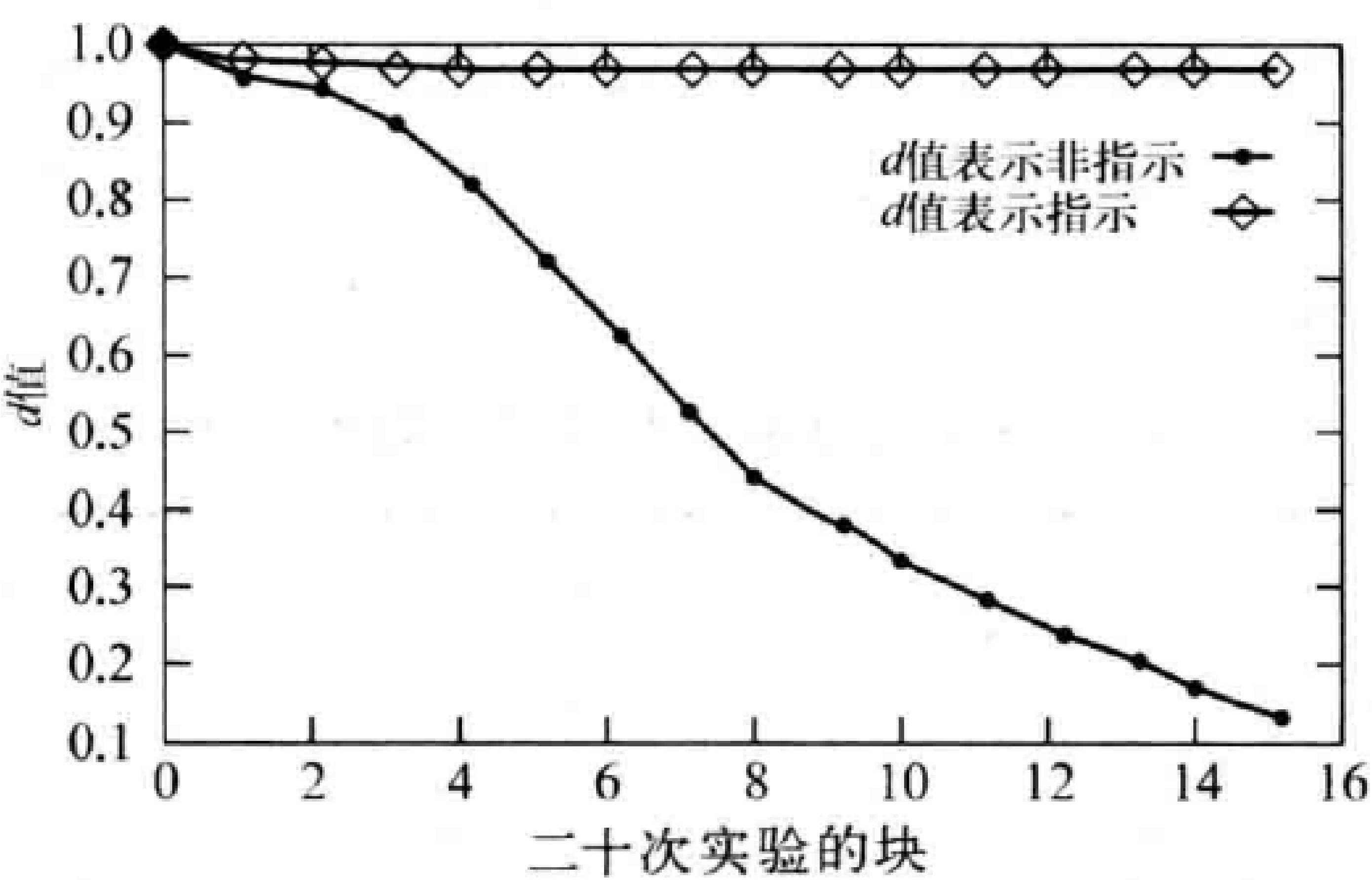


图 4　英语中六十个物理应用题的平均指示学习曲线

类 $\gamma3$ 对于法语来说是特殊的,因为要求用几个词把 *medium* 的思想表达在 *medium screw* 中,即短语 *de taille moyenne*。我们承认,最终分析是对法语的自然语义的一种曲解。这种曲解源于英语词汇的偏见,而不是内在于我们的内部语言中指示词的固定集合的文法偏见。

假定汉语不是屈折语(inflected language),德语是屈折语,表 4 中的非指示词数很大。我们无法详细讨论所有这些虚词。虚词 *ba* 出现在描述直接行动对象的名词短语之前。虚词 *zai* 出现在描述关系对象的名词短语之前。我们注意到,表 5 中给出的汉语文法有下列两个规则:

$$A \rightarrow [\gamma8]\ O\ A_3[\gamma11]\ G$$

和

$$G \rightarrow REL\ [\gamma6]\ O$$

但是,把 *zai* 放在一个关系词之前是不正确的汉语。这个超集就理解而言是不成问题的,正确的例子是,运用 ϵ ,而不是 *zai*。非常类似于对 *zai* 的那些评论,也适用于 *chao*。在它们的用法上的不同反映了一个语义上的区别:对理解而言我们不需要,但对产生式而言我们需要。虚词 *zai* 的用法主要与像 *on* 和 *under* 之类的位置关系相关,相反,*chao* 的用法与运动方向相关。

文法规则。表 5 中给出了十种语言的最终文法的概述。这

439

表 5　十种可比较语料库的理解性文法

		E	D	G	F	S	C	R	Ch	K	J
Ⅰ	$A \rightarrow A_1 + O$	+	+	+	+	+	+	+			
	$A \rightarrow [\gamma8] + O + [\gamma9] + A_1$								+	+	+
Ⅱ	$A \rightarrow A_2 + [\gamma14] + G$	+	+	+	+	+	+	+			
	$A \rightarrow [\gamma12] + G + [\gamma13] + A_2$								+	+	+

(续表)

		E	D	G	F	S	C	R	Ch	K	J
Ⅲ	$A \rightarrow A_3+O+G$	+	+	+	+	+	+	+			
	$A \rightarrow [\gamma 8]+O+A_3+[\gamma 11]+G$								+		
	$A \rightarrow O+[\gamma 9]+G+[\gamma 10]+A_3$									+	+
Ⅳ	$A \rightarrow A_5+DIR+O$	+									
	$A \rightarrow A_5+O+DIR$		+								
	$A \rightarrow DIR+O+A_5$			+							
	$A \rightarrow [\gamma 8]+O+[\gamma 9]+A_5+DIR$								+	+	
	$A \rightarrow DIR+A_5+O$				+	+	+	+			
	$A \rightarrow O+[\gamma 7]+DIR+A_5$										+
Ⅴ	$A \rightarrow A_5+O$	+	+	+	+			+	+		
	$A \rightarrow O+[\gamma 7]+A_5$										+
Ⅵ	$G \rightarrow REL+[\gamma 6]+O$	+	+	+	+	+	+	+	+		
	$G \rightarrow O+[\gamma 5]+REL$								+	+	+
Ⅶ	$DIR \rightarrow REL$	+	+	+	+	+	+	+	+	+	+
Ⅷ	$O \rightarrow [\gamma 2]+S$	+	+	+	+	+	+	+	+	+	+
	$O \rightarrow [\gamma 7]+S$										+
Ⅸ	$O \rightarrow [\gamma 1]+S$	+	+	+	+	+	+	+	+	+	+
	$O \rightarrow [\gamma 7]+S$										+
Ⅹ	$S \rightarrow PROP+[\gamma 4]+S$	+	+	+	+			+	+	+	+
	$S \rightarrow S+[\gamma 3]+PROP$				+	+	+				
Ⅺ	$S \rightarrow OBJ$	+	+	+	+	+	+	+	+	+	+

个表有 11 列,在第一列,我们以内部语言文法规则的顺序(参见表 1)列出了文法规则。下一个列,我们以下列顺序列出语言:英语、荷兰语、德语、法语、西班牙语、加泰罗尼亚语、俄语、汉语、韩语、日语。每一个内部文法规则通常都对应于一个语言特殊规则的集合。在规则Ⅶ和Ⅺ的情况下,这个集合是一个单元集合,因为显而易见的理由是,在我们有限的内部语言中不可能发 440

生任何变化。显然,在更一般的情况下,例如,在规则Ⅶ的情况下,关系可能有修改特性。

在印欧语系和亚洲语系之间最重要的对比是,在印欧语系中,表达行动的使役动词通常在一种表达的开始,但在亚洲语系中,通常放在最后。比如,参见从规则Ⅱ派生出的两个规则。第一个规则针对七种印欧语言,第二个规则针对三种亚洲语言。类似的评论对于从规则Ⅲ导出的三个规则也成立。

在这两组语言之间的这种众所周知的对比,带来了关于表5中给出的规则的更系统的问题。与内部语言的每个规则相对应的规则集,穷尽了语义范畴顺序的可能排列了吗?令人惊讶的是,除了从规则Ⅲ生成的集合(这个集合只有三项而不是六项)之外,答案是肯定的。

然而,只对我的两位同事和我所掌握的三种母语进行反思,我们就能在词汇表中和在显示出缺少两种排列的我们的不同语料库的概念框架内,举出一个简单的例子:

	A_3	G			O	
E	*Put*	*near*	*the*	*washer*	*a*	*screw.*
G	*Leg*	*neben*	*die*	*Scheibe*	*eine*	*Schraube.*

	G			A_3	O	
E	*Near*	*the*	*washer*	*put*	*a*	*screw.*
G	*Neben*	*die*	*Scheibe*	*leg*	*eine*	*Schraube*
Ch	*Zai*	*nage*	*dianquan*	*fujin fang*	*yige*	*luosiding.*

至于第三种缺少的排列,我们的韩语合作者为我们举出了下列例子:

	G			O		A_3
K	*Ku*	*nasapati*	*yephey*	(han)	*nasa-lul*	*nohala.*
	The	washer	near	(one)	screw	put.

作为原则性问题,表1和5基本一致是误导性的。尽管表5的文法规则是通过公理 *W*3 直接从内部语言的原始文法规则中派生的,正如表1中给出的那样,但情况未必如此。更大的语料库有要求在应用公理 *W*3 时派生内部文法规则的事例,这个语料库是这里针对英语、汉语和德语所研究的语料库的一个超集。这种现象的一个德语例子取自苏佩斯、伯特纳和良等人的文章(1995),这个例子是,

$$A \rightarrow A_4[einen]\ S\ D\ [der]\ PROP\ ist。$$

这个规则的一个实例是,

Heb eine Mutter hoch, die nicht groß ist。

相关工作和未解决的问题。到目前为止,对语言学习的最广泛研究是关于儿童学习母语的研究。很详尽的理论论述可在韦克斯勒(Wexler)与库利卡沃(Culicover)的著作(1980)和平克(Pinker)的著作(1984,1989)中找到。这个工作有些与我们的工作相关,但在理论思想细节的核心方面并不相关。例如,韦克斯勒和库利卡沃假设,学习者已经有所学语言的成熟的语境无关的文法,而且,学习主要是集中于学习转化。相反,我们开始时没有所学语言的文法。我们没有详细地阐述我们的工作与儿童学习的相关性,但在某种程度上,联系无疑是存在的。 441

在过去的十年左右,在自然语言的机器学习方面的文章或著作相对较少。兰利(Langley)和卡博内尔(Carbonell)的文章(1987)提供了直到他们的文章刊出前的一个出色的概述。在这项工作的框架内,我们的工作受到的语义驱动大于句法驱动。费尔德曼(Feldman)等人的工作(1990,1996)和西斯金德(Siskind)的工作(1992,1994)也分享了这种语义承诺,他们的研究也最接近于我们自己的研究。费尔德曼等人用直接简单的

术语描述了他们最初的想法(1990)。首先,向学习系统呈现出两幅图和关于这两幅图的真自然语言的陈述。第二,这个系统要很好地学会这种语言,足以确定一个新的句子是否对附图来说是正确的。费尔德曼等人的语言学习进路是,把文法学习与词汇概念的学习分离开来(1996)。运用在可能文法的集合上的贝叶斯推理和模型合并学习文法。西斯金德当初的工作(他的学位论文)是在朴素的物理学语境中进行的,但也关注儿童在语言学习时可能运用的算法(1992)。西斯金德在1994年的文章中,还在继续这项工作,不过,淘汰了前语言知识的任何假设。把注意力集中于经过含义的可能内部表征获得词汇。尽管西斯金德(1994)完全集中于词汇的含义,但是,他的构成一种学习算法的七步程序,在最高水平上,而不是在细节上,与我们的程序相似。西斯金德提出了一个概念(1991),这个概念与我们在非指示词的极限条件下等于0的指示值的概念肯定不同,但在某些方面相类似。然而,他的思想不是概率的,而且,他没有提供任何学习曲线。在他的温度概念中,他确实提供了对同音异义的一种论述(1994),这是我们没有的。

尽管我们考虑十种语言,但当前对我们理论的检验,在特征上,一定会被认为是很初步的。我们有必要把这个理论推广到解决各种迫切需要解决的问题。我们限于四个问题,但在这个领域内(即我们的工作与其他人的工作中),这是理论发展的特征还很初步的一个标志,任何一位见多识广的读者能够很快地把这个名单翻两三倍。在我们的观点中,对于我们正在发展的这种理论而言,直到解决了更多的概念问题时,大规模的实验才是成熟的:

指代学习。即使很有限的和有时完全是人为使用的自然语言,通常也充满了指代的用法。物理应用题就是一个好的例子。

学习时间特征。语言的日常用法以丰富多样的方式标出了事件的时间序列。科学技术中的许多系统论述要求连续时间和必须学会的时态区别。 442

学习多重含义。在机器人的用法中已经有一些棘手的例子,例如,screw 在英语中既是名词又是动词。更加奇特的例子是 washer 在英语中有许多含义:金属环(我们的机器人的案例),洗衣机,浣熊,等等。

学习概念和词。可能更好的是主张,应该学习概念,而不只是学习代表概念的词。尽管我们已经开始了包括多变量的网络模型在内的某种系统工作,但在当前,我们的理论在这个方面还是无可奉告(Suppes and Liang, 1998)。

关于学习的最终评论。联想的一般思想或条件作用(这种机制在 3～5 节中是学习的核心)的重要性是不会降低的。现在,在当前的神经科学的语境中,一种主要挑战是找到它的物理体现。为了达到这个目的,在进化树上海兔以上动物的许多经典条件作用实验,对于鉴别何时何地发生条件作用来说,是必需的。追溯关于行为主义的修辞的哲学家通常意识不到,当在第 3 节和第 5 节中对条件作用理论进行公理化的阐述时,联想概念、条件作用状态乃至刺激概念都有主要的理论地位,不是直接的可观察量,只可能从行为反应数据中间接地估计。

另一方面,似乎明显的是,用非常特殊的机制表示不同类型的学习。这种情况几乎是毫无疑问的:例如,学会一个新词,必须通过极其大量的神经元,绝不只通过一个神经元。这种学习很可能也是采取了这样的学习形式:用给定振幅和相位的学习频率代表大脑皮层中的词。对于其他的大脑活动而言,也能为通过单个神经元的学习想象不同的表征,正如现在的许多重要例子中证实的那样,但不适合于人类语言,我们现在转

向这个主题。

8.6 语言与大脑

某种历史背景。 亚里士多德说过,人作为动物具有的显著特征是,人是理性的动物,但是,用更加生物学和心理学的术语来说,这个特征是,人是说话的动物。在我们还没有完全探索的那些方面,语言是作为动物的人的最显著的标记。不仅对于言语的产生式来说,而且对于有意形成说什么或对所见所闻的理解来说,语言的处理首先集中在大脑中。因此,语言的大脑处理是本节的焦点。我首先对大脑中电活动的发现作出历史性的概述。

关于通过动物的肌肉或神经产生电的一个早期参考文献来自弗朗切斯科 · 雷迪(Francesco Redi, 1671)的一项研究,他描述了这方面他在1666年做过的一个实验,"似乎对我来说,好像电鳐的痛苦反应,与任何其他部位相比,更多地位于两块镰形部位或肌肉中。"雷迪的工作是在梅迪奇(Medici)的指导下,在佛罗伦萨完成的。这些电观察是片断的和不充分的。然而,各种动物的肌肉或神经中的电活动思想成为整个18世纪的潮流(Whittaker, 1910)。在雷迪之后和路易吉 · 加尔瓦尼(Luigi Galvani)在博洛尼亚迈出决定性的一步之前,还间隔了一百多年。加尔瓦尼以下列方式描述了他的重要步骤:

443

> 以下列方式工作进程已经取得了进步。我解剖一只青蛙并准备好后……出于其他考虑,我把这只青蛙与一台电子仪器放在同一张桌子上。当我的一位助手偶尔轻轻地用解剖刀尖碰到青蛙腿部内侧的神经时,看到所有肢体的肌肉突然收缩,似乎陷入了猛烈的强直性惊厥。当我们正在

做电实验时，另一位在场的助手曾认为，他看到在电子仪器的导体喷射出一个电火花时，发生了这种现象。这很令人诧异，他立即提醒我注意这个不寻常的现象，当时，我正在全神贯注于其他问题。于是，我以极大的热情渴望重复这个实验，以便澄清这种模糊现象并使其为人所知。因此，在某一位助手产生了一个电火花的同时，我自己用解剖刀尖依次接触腿部的神经；这种现象本身以和从前完全一样的方式重复发生。

(Galvani，1791/1953：pp. 45－46)

加尔瓦尼 1791 年的工作受到著名的意大利物理学家伏打(Alessandro Volta)的强烈批评，伏打出生在科摩，是帕维亚大学的一名物理学教授。下面是在伦敦皇家学会读到的他的一段批评，摘自伏打写给卡瓦洛(Tiberius Cavallo)的信。

动物电的名称，在加尔瓦尼等人意指的意义上，一点也不恰当；也就是说，动物器官本身固有的力，即生命力的某种特殊作用，使得动物器官中的电流体不平衡。不，这只不过是一种外因导致的人为的电，也就是说，最初通过金属与任何一种湿的物质的连接以目前未知的方式激发的。因此，动物器官，即神经和肌肉，只不过是被动的，虽然只要它们，特别是神经，被置于以上提到的方式产生的电流回路中，受到电流的攻击和刺激，就很容易起作用。

(Volta，1793/1918：pp. 203－208)

加尔瓦尼能够直接应对这些批评，他在 1794 年发表了一篇匿名回应，其中包括对一个不用金属的肌肉收缩实验的详细说明(Galvani，1794)。加尔瓦尼工作的原创性和重要性在整个欧洲得到认可。杰出的德国物理学家埃米尔·杜波依

斯-雷蒙德(Emil Du Bois-Reymond)以下列方式总结了加尔瓦尼的贡献:

> 1. 动物具有它们自己特有的一种电,称之为动物电。
>
> 2. 与这种动物电最密切的器官和分布动物电的器官是神经,动物电的最重要的分泌器官是大脑。
>
> 3. 神经的内物质是专门用来导电的,而外脂肪层阻止这种物质的分散,允许它的积聚。
>
> 4. 动物电的接受器是肌肉,而且,这些肌肉类似于外侧带负电内侧带正电的一个莱顿瓶。
>
> 444 5. 运动机制在于肌液的放电,从肌肉内部经过神经到达外部,而且,肌肉莱顿瓶的这种放电向敏感的肌肉纤维提供了电刺激,因此引起肌肉收缩。
>
> (Du Bois-Reymond, 1848/1936: p. 159)

下一个重要的事件是卡罗·马泰乌奇(Carlo Matteucci)证明:电流源于肌肉组织。然而,大约在加尔瓦尼之后一百年,利物浦的理查德·卡顿(Richard Caton, 1875)用汤姆孙[Thomson,开尔文勋爵(Lord Kelvin)]的反射电报式电流计在外露的兔脑中检测到电活动。1890年,波兰的阿道夫·贝克(Adolf Beck)在狗和兔的大脑皮层检测到规则的电图形。到了19世纪末,一位荷兰医生和生理学家艾因特霍芬(Willem Einthoven)①基于他以前发明的所谓弦线电流计,研制出了一台新的心电图仪,这台仪器类似于探测穿越大西洋电缆的电报信号所开发的装置。由于艾因特霍芬的弦线电流计具有很高的

---

① 原文是"Villem Einthoven"这个名字有误,应该是"Willem Einthoven",本文经与作者沟通后更正。——译者

灵敏度，所以，1914 年，波兰克拉科夫大学的齐布尔斯基(Napoleon Cybulsky)和马切斯纳(S. Jelenska Macieszyna)用它记录了一只狗的癫痫发作。大约从 1910 年开始，汉斯·伯格(Hans Berger)在德国耶拿开始了透过完好的头颅检测电活动的大量的系列研究。这对应用于人类具有很大的意义。他的观察于 1929 年发表，但完全没有得到大家的认可。然而，直到他的发现被剑桥大学的爱德华·道格拉斯·阿德里安(Edward Douglas Adrian)和马修(B. H. C. Matthews)确证之后，才开始得到大家的赏识，他们在 1934 年的剑桥生理学学会上和 1937 年的国际心理学大会上证实了伯格的发现。20 世纪 30 年代末和 40 年代初，关于大脑中电活动的研究，或者，我们现在所取名的脑电图(EEG)首先进入北美洲——哈佛医学院的伦诺克斯(W. G. Lennox)、埃尔纳(Erna)和吉布斯(F. A. Gibbs)，布朗大学的唐纳德·林斯利(Donald Linsley)和麦吉尔大学的怀尔德·彭菲尔德(Wilder Penfield)。最初证实伯格工作的英语报告之一是由贾斯帕(Jasper)和卡迈克尔(Carmichael)完成的。几乎同时，吉布斯等人(1935)在美国开始使用由加尔索(Garceau)和戴维斯(Davis)制作的(1935)第一台用墨水书写的电报式 EEG 自动记录仪。[①]到 20 世纪 50 年代，EEG 被广泛地应用于临床，特别是应用于癫痫的研究和大脑中电活动本性的各种研究。本书没有详细地总结从 1950 年到现在的很多科学家的工作，但是，关于 EEG 的一个出色评论，也就是说，与认知特别相关的电活动的一个出色评论，可以在鲁格(Rugg)和科莱斯(Coles)的著作(1995)中找到。

观察大脑的活动。当前观察大脑活动的四种主要方法是很

---

① 我从格迪斯(Geddes)的文章(2001)中获得了这些最新的参考文献。

容易描述的。第一种方法是已经提到的经典的脑电图(EEG)观察,这一点很重要,它至少有千分之一秒的时间分辨率。第二种方法是现代的磁场观察,而不是电场观察,这是在脑磁图(MEG)的标题下进行的。这种方法也有近似于千分之一秒的相同的时间分辨率。第三种方法是正电子成像术(PET),这在最近几十年里得到了广泛的使用,在某种情况下,这种方法有利于观察大脑活动的定位,但只有一秒的时间分辨率。最后,当前最流行的方法是功能性磁共振成像(fMRI),它的最大优点是,在大脑中很好定位的地方,观察能量的吸收,但不幸的是,也有不小于一秒的时间分辨率。

445

尽管通过PET和fMRI能了解到许多出色的图像,但是,如果人们想识别代表词或语句的脑电图,它们实际上是无用的,因为处理过程(尽管根据现代计算机的标准看是很慢的)太快啦,以至于用几乎等于一秒的观察时间分辨率会一事无成。例如,典型的词,不管是听还是读,都在四百或五百毫秒之内甚至通常将更快得到处理。我自己真正的兴趣,即关注大脑处理语言的方式,是在1996年我听了萨姆·威廉森(Sam Williamson)关于MEG的一场精彩报告后受到的激励开始的,萨姆·威廉森在MEG的发展中从一开始就是一位杰出的物理学家。我怀疑他所说的内容,但是,关于这一点,我越想越意识到,试图用MEG辨识单个词的处理过程是有趣的和重要的。这种想法使我联想到了本质上类似于语音识别的作为MEG记录的一个脑电波识别程序。我很熟悉从20世纪40年代到现在的语音识别的长期历史,而且,我认为,也许同样强烈的分析努力能够产生与相应结果一样的某种东西。因此,在1996年,特别是在组约大学跟随萨姆·威廉森和劳埃德·考夫曼(Lloyd Kaufman)刚获得博士学位的吕忠林

(Zhong-Lin Lu)[1]协助下，我们在加利福尼亚圣地亚哥的斯克里普斯研究所做 MEG 实验。当我们分析第一个实验结果时，根据听到的一个词或从计算机屏幕上读出的一个词，识别正在处理七个词中的哪一个词的问题，我们不能从 MEG 的记录中得到很好的识别结果。幸运的是，在斯克里普斯的 MEG 设备(它在技术上比运行标准的 EEG 设备更昂贵和更复杂)上进行的实验，也记录了多年来用标准的二十台 EEG 传感器的数据。我们也着手分析 EEG 的数据，在这里，我们有许多更好的识别结果(Suppes, Lu and Han 1997)。

在标准的 EEG 系统中(在观察大脑中的电活动时，全世界广泛地使用的一个系统)，记录电活动的传感器在通常所谓的10～20 系统中进行排列，如图 5 所示，定位在一个近似于圆形所示的头颅表面，耳朵在图的左右两边，眼睛在图的上部。首先，在这些位置上所用首字母对应于大脑部分的参考定位，例如，F 表示额(frontal)，C 表示中枢(center)，T 表示颞(temporal)，P 表示头顶骨(parietal)，O 表示枕骨(occipital)。其次，你会注意到，单号传感器定位于头颅的左半球，双号传感器定位于头颅的右半球，三个传感器大约沿中枢线定位。在语言处理的过程中，左半球或右半球可能两者同时发生了什么，关于这个问题，现有的观点多于能深入阐述和确立的事实。我自

---

[1] 吕忠林，美籍华人，现在为美国南加州大学 William M. Keck 认知神经科学讲座主任，心理学和生物医学工程教授，Dana and David Dornsife 认知神经科学成像中心主任。他在南加州大学组建了脑处理实验室，主要研究目标是建立感知与认知的数学模型。吕忠林的研究领域主要包括：(1) 视觉与听觉认知，选择性注意，知觉学习的计算和心理物理研究；(2) 知觉，选择性注意，语言与决策的脑功能成像研究；(3) 阅读障碍、弱视与老年痴呆的视觉损伤。吕忠林于 2002 年被选为中国科学院海外评审专家，于 2003 年获得美国实验心理学会杰出青年奖，并成为该学会最年轻的高级会员。——译者

已的观点是也许存在着比一般认可的观点更多的二元性,但在当前的语境中,我将不打算对这种观点作出一种经验的辩护,尽管我已经发表了支持二元性的数据(Suppes, Han and Lu 1998,表 2)。

446

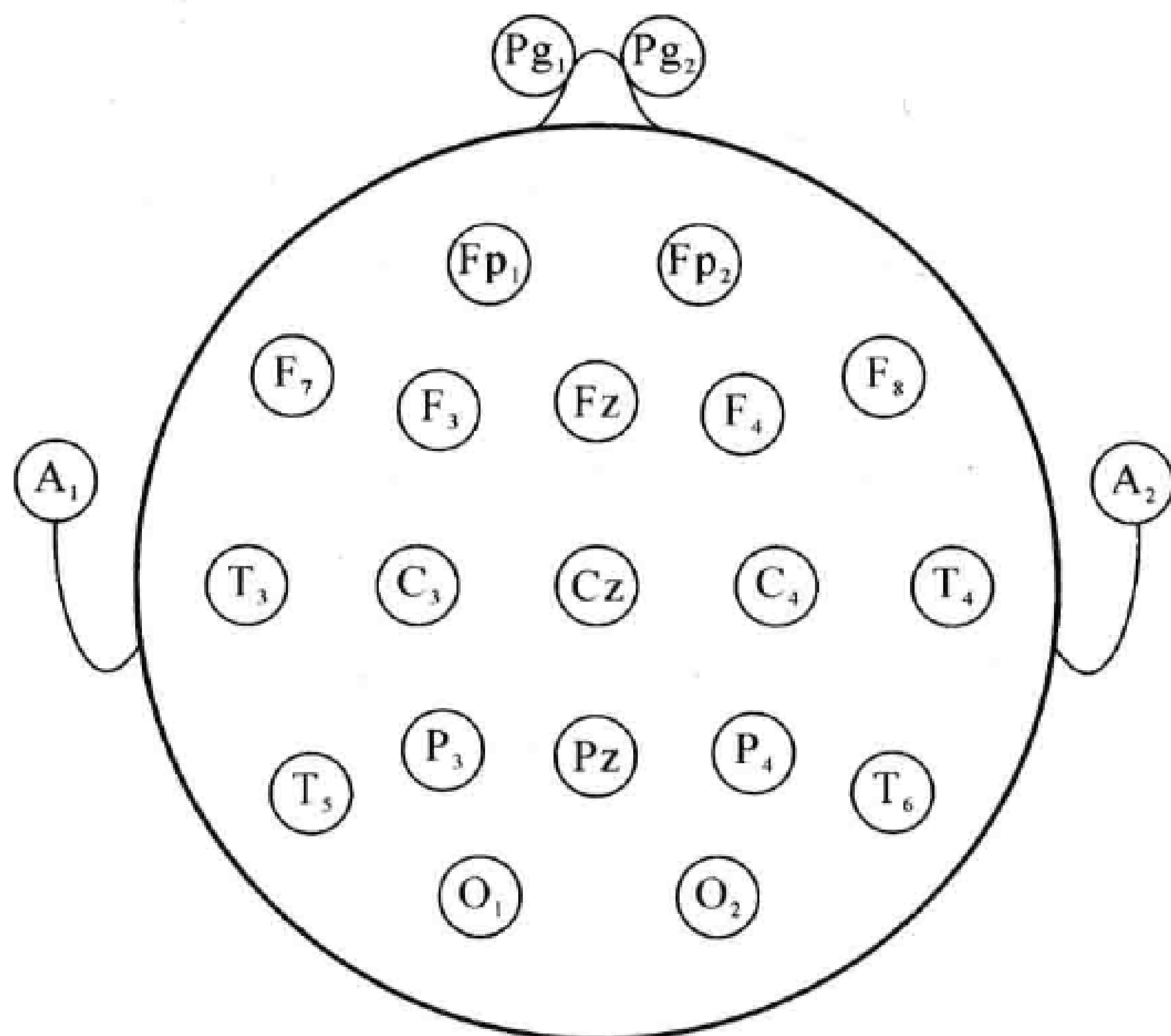


图 5 EEG 传感器的 10～20 系统

一个典型的试验如图 6 所示:在这个试验中,向受试者提供一个能看见的句子,即在一台计算机屏幕上逐词打出来的一个句子。试验持续三千多毫秒,把每毫秒观察到的波形的一个振幅以微伏为单位标在 Y 坐标上。已知在三秒多一点的时间内只从二十台传感器中的某一台传感器观察到如此多的数据,由此不难看出,语言活动的 EEG 记录的数据很丰富,事实上,我们几乎可以说被数据淹没了。对几位受试者做一个实验,是不困难的,每一位受试者接受几个小时的记录,最终得到如图 6 所示的在五吉和十吉字节之间的那类数据。这意味着,企图发现对应于特殊词和语句的波形的问题不是一件简单之事。它完全不

同于研究语言的行为实验,在这个实验中,在几个兆字节内很容易记录所观察到的受试者的反应,然后没有如此大量的计算而作出分析。

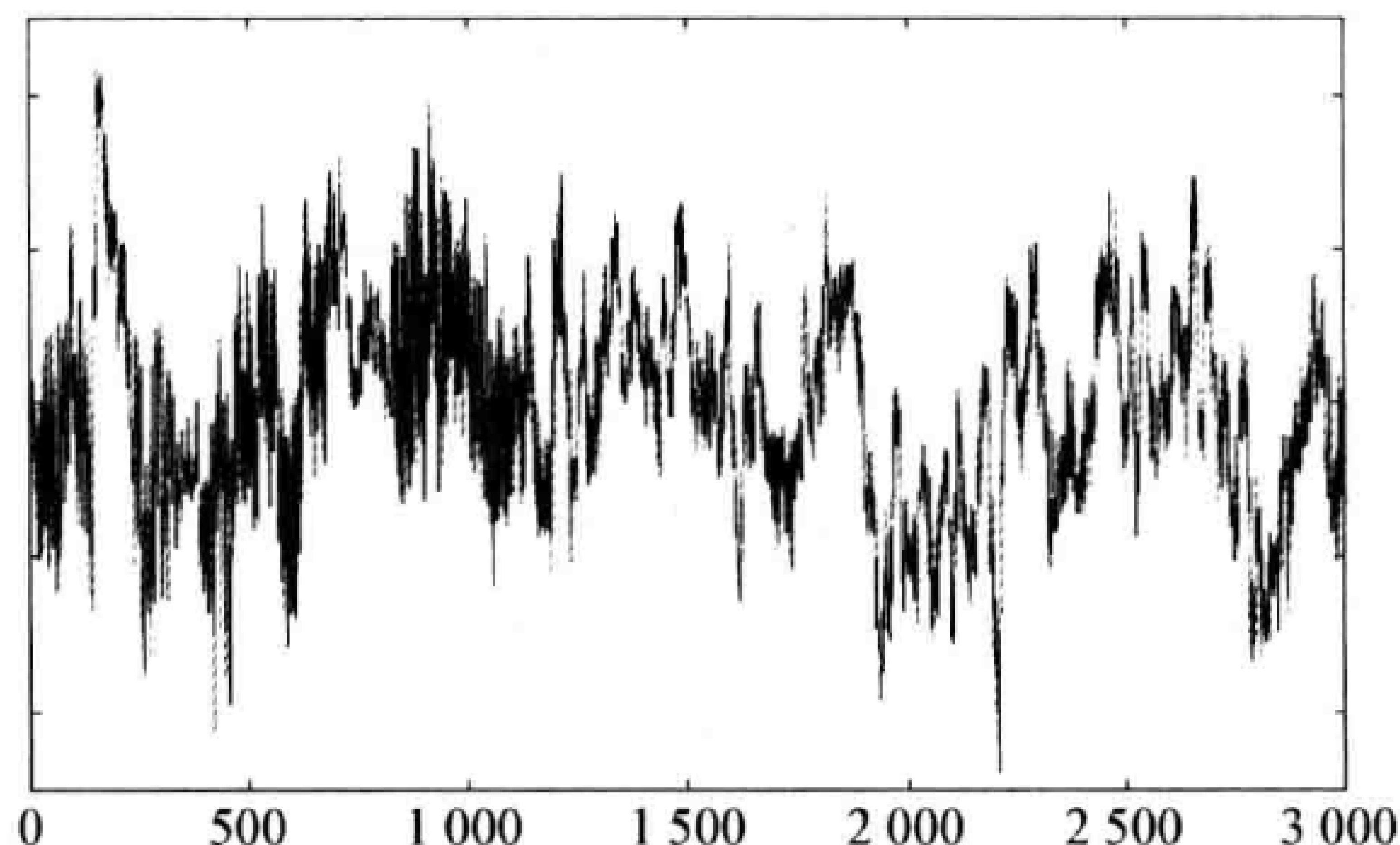


图 6　从一个传感器和一个试验获得的原始 EEG 样本数据

数据分析方法。在刚才描述的那类数据中找出词和句子没有捷径可走,因此我在这里略述我和我的同事在过去几年内在某种程度上成功地运用的方法。基本进路在很大程度上采取了数字信号处理,但这种应用与电气工程师或在数字信号处理方面运用五十年前提出的大量数学、定量和统计技巧的其他人通常所关注的应用完全不同。一个极好的一般参考文献是奥本海姆(Oppenheim)和谢弗(Schafer)的著作(1975)。

这条进路是容易概述的,尽管技术细节更复杂,只能略述。一般情况下,除了注意到的一些重要例外,要做的第一件事情是对某个给定条件下的数据求平均,例如,一种典型的情况是,在给定的听力或视力条件下,大脑对一个特殊的言语刺激的反应。根据这种信号的频率较低的假设,求平均的目的是消除噪音,特别是高频噪声。然后下一步是,根据平均数据进行傅里叶变换,从时域转换到频域,也许这是信号处理分析的最典型的特征。

第三步是在频域内进行筛选，更进一步降低鉴别所处理的词或语句所用的信号的带宽，这一步能够与第二步同时进行。我们在某些工作中相当彻底地探索的一条替代进路是，在频域内根据计算其振幅的绝对值挑选具有最高能量的频率，然后，在时域内叠加这些正弦函数。[①]当我们经过一个傅里叶逆变换回到时域时，不管是通过滤波，还是通过叠加，在任何一种情况下，我们都获得了一个更加简单的信号。

现在暂时忽略叠加方法，只考虑滤波，我们返回到一个带通滤波器的时域，这个带通滤波器是由低截止频 L 和高截止频 H 这两个参数确定的。我们现在挑选另外两个重要参数。也就是说，在给定的实验条件下，我们应该如何选择词或句子的信号的开始 s 和结束 e。像在其他脑问题中一样，在观察到的连续活动的电波中，乍一看所记录的脑电波，一个特殊的词或句子的表征何时开始或结束，是不明显的。当选择出一个四元组(L, H, s, e)时，我们就用这个四元组参数对在给定的实验条件下作为刺激源的词或句子的脑电波进行分类。

448

我们的目标是使正确地归类的脑电波的数量最优化或最大化。我们一直调整这个四元组参数，直到获得能够找到的似乎是最佳的结果时为止。我们在一个四维的参数网格中进行优化，不过，我在图 7 中表明了这个滤波器的两个参数的最佳平面，尽管这里把高截止频率和低截止频率之差，而不是高截止频率本身，用作纵坐标或 Y 轴的计量单位。因此我们有，横坐标是以赫兹为计量单位的低截止频率，纵坐标以频率之差 W 为计量单位，也就是以赫兹为单位的带通滤波器的带宽。图 7 所示

---

① 实际上，在"挑选振幅为 A_i 的频率 ω_i"时，我们事实上是在挑选一个正弦波，$A_i \sin(\omega_i t+\varphi_i)$，这里，$\varphi_i$ 是它的相位。下一页更多地讨论这一点。

的这个平面的平滑度是我们通常所观察的颅外(即我们正在观察这些观察值的地方)电场的特征,也是预期的电场观察的特征。(我们当然非常幸运,电场很强,在完全不受噪音干扰的条件下足以能被观察到。)这个等高线图代表了下面更详细讨论的用四十八个地理学的句子做的第一个实验,但这个图清楚地表明,最好的识别率是四十八个句子中识别出四十三个(大约90%)。当带通滤波器的参数发生变化时,等高线图一定表现出低识别率。

449

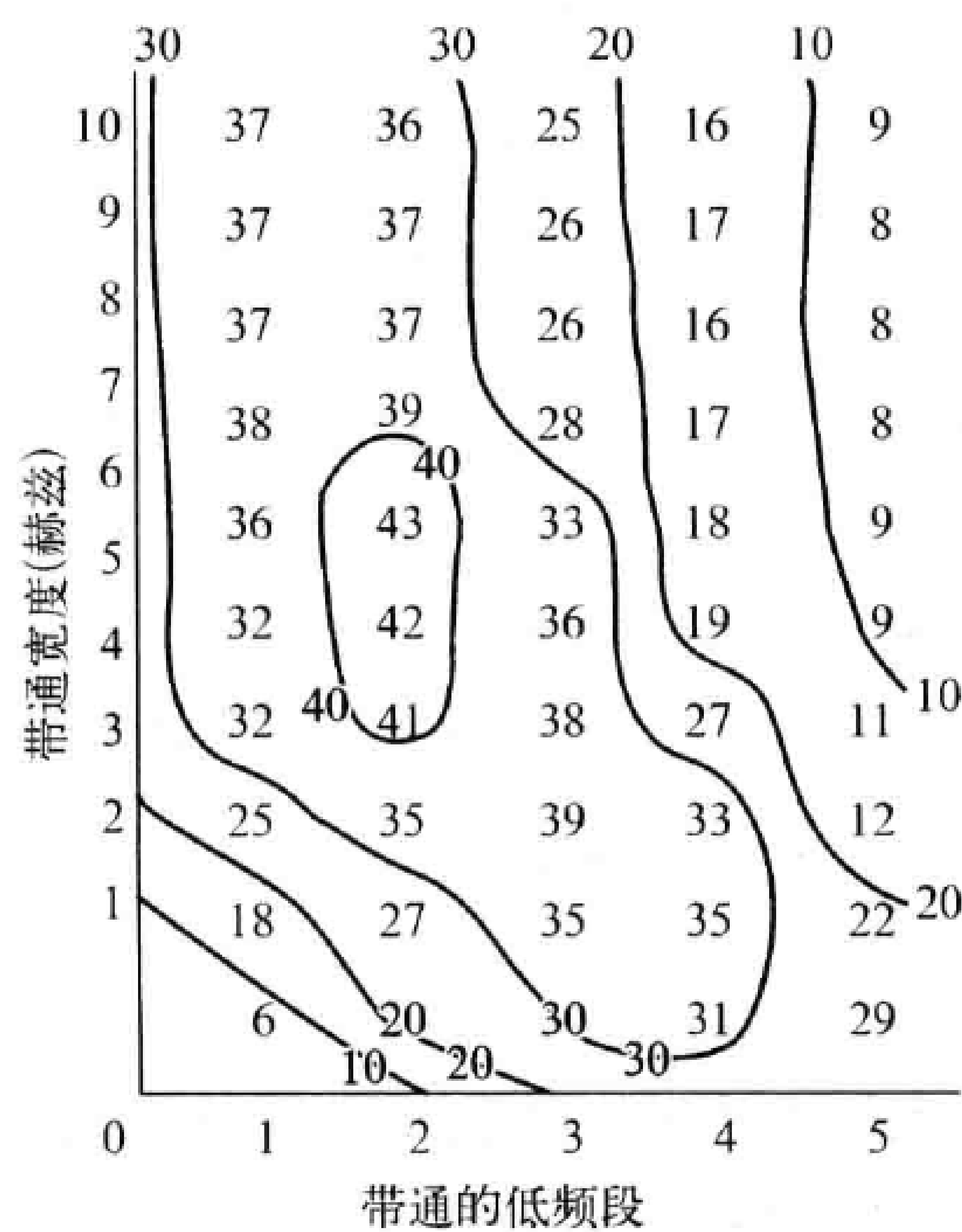


图 7　带通滤波器参数 *L* 和 *W* 的识别率平面的典型等值线图

EEG 数据的傅里叶分析。傅里叶分析的背景是标准的傅里叶积分理论,但在实践中,我们的数据是有穷的。我们有兴趣观察的有穷脉冲数据通常只持续几秒。例如,所研究的句子,在以自然的速率讲出时,通常持续不超过三秒或四秒。我们运用

离散傅里叶变换分析具有给定振幅和相位的频率。正如所表明的那样,我们的目标是找到含有信号和消除噪声的频率。(本节最后讨论由眨眼或其他这样的事件所产生的伪差。)

设 N 是观察值的个数,观察的时间间隔相等,通常为 1 毫秒。我们于是用傅里叶级数的系数 $\tilde{X}(k)$ 把观察值 $x(n)$, $0\leqslant n\leqslant N-1$ 的有穷序列表示为周期为 N 的周期序列,因此,我们有对偶对

$$\tilde{X}(k)=\sum_{n=0}^{N-1}\tilde{x}(n)e^{-i\left(\frac{2\pi}{N}\right)kn} \tag{1}$$

$$\tilde{x}(n)=\frac{1}{N}\sum_{k=0}^{N-1}\tilde{X}(k)e^{i\left(\frac{2\pi}{N}\right)kn}。 \tag{2}$$

我首先注意到下列内容:

1. 周期性的正弦波和余弦波用标准的指数项来表示。

2. $\tilde{x}(n)=x(n)=\tilde{x}(n+kN)$,波浪字符~表示长度 N 的周期性,也代表时间和频率的对偶性。

3. 指数式中的 kn 向我们提供了独特的指数,因此,正弦项和余弦项代表周期 N 的整数约数。我们用这种方法表示$\frac{2\pi}{N}$的整数倍数的频率。

4. 运用周期性 N 使我们得到了时域与频域之间的对偶性。

这两个方程(1)和(2)的特性是

1. 线性:如果 $\tilde{x}_1(n)$ 和 $\tilde{x}_2(n)$ 有周期 N,则

$$\tilde{x}_3(n)=a\tilde{x}_1(n)+b\tilde{x}_2(n)$$

450 和

$$\tilde{X}_3(k)=a\tilde{X}_1(k)+b\tilde{X}_2(k)$$

也有周期 N。

2. 在序列移位条件下的不变性

$$n \to n + m$$

3. 各种对称性,例如,$|\tilde{X}(k)| = |\tilde{X}(-k)|$

4. 周期为 N 的 $\tilde{x}_1$ 和 $\tilde{x}_2$ 的卷积有周期 N:

$$\tilde{x}_3(n) = \sum_{m=0}^{N-1} \tilde{x}_1(m)\tilde{x}_2(n-m)$$

重要的是有效的快速离散傅里叶变换,这是由库利(Cooley)和图基(Tukey,1965)等人发明的一种算法,下面的计算中用了这种算法的一个变种。

滤波器。滤波器构造的原理很简单。细节却并非如此。一个带通滤波器,例如,1～20 赫兹滤波器只"过滤低于 1 赫兹和高于 20 赫兹的所有频率"。关于设计滤波器的工艺和理论的电气工程文献有许多进展,这里不可能对此作出概述。重点始终是设计出能达到某种最优化标准的一个滤波器。

如果信号是已知的,那么工程目标就是使信号的传送达到最优。正如已经提到的那样,我们的问题是,在我们的实验中,携带所研究的词或句子信息的信号是未知的。因此,我们的解决办法是,使预测正确分类的滤波器达到最优。我们所用的参数上面已经讨论过。此外,我们通常通过运用一个四阶巴特沃思滤波器对滤波器的边缘进行平滑校正,尽管在这里汇报的工作中,更简单的方法也恰好能达到这个目的。[①]

三种实验结果。我现在转向我们到目前为止所获得的最重要的三种结果。

受试者之间的不变性。在第一个实验中,我们向九位受试

① 奥本海姆和谢弗的著作中很好地描述了巴特沃思滤波器(1975,pp. 211－218)。

者提供四十八个关于欧洲地理的句子。要求受试者判断这些句子的真假，同时，当他们听到这些句子或读出在一台计算机屏幕上逐词显示的这些句子时，我们进行典型的EEG记录。语义学的任务是简单的，但因为这些句子间只相隔四秒，所以，判断这些句子真假的任务并不是小事。典型的句子形式是：**意大利的首都不是巴黎，华沙是奥地利最大的城市**。现在，记下来自五位受试者的数据，通过对五位受试者的数据求平均，形成这四十八个句子的原型，而且，记下来自其他四位受试者的数据，形成每个句子相应的平均后的测试样本，我们应用上面描述的傅里叶方法，并从预测的立场出发找到了最佳带通滤波器。[①]滤波之451 后，把四十八个原型中每个原型与一个测试样本之间的经典最小二乘法拟合，用作选择标准，我们对试测样本的正确识别率能达到90%(Suppes, Han, Epelboim and Lu 1999a)。[②]

令人惊奇的不变性结果是，原形数据和测试样本数据来自不同的受试者。这两个小组间的人员是不重叠的。从理论上看，这是用于通讯系统的一个有效方面。我的词或句子的脑电图表征与你的表征很类似，因此，理解你是容易的。但是，一种理论观点要得到人们的接受，需要有很强的经验支持。另一个评论的角度是，当不同受试者完成相同的任务时，他们大脑皮层

---

① 数据表示读出这些句子的视觉条件，一次只显示一个词。

② 设$x_i(n)$，$0 \leqslant n \leqslant N-1$是第$i$个原型(在时域内)和$y_j(n)$，$0 \leqslant n \leqslant N-1$是第$j$个测试样本。于是，平方差之和是$S_{ij}$，这里

$$S_{ij} = \sum_{n=0}^{N-1} (x_i(n) - y_j(n))^2$$

如果

$$S_{jj} = \min_i S_{ij}$$

且最小值是惟一的，则对测试样本$y_j(n)$的分类是正确的。

里的电活动比他们大脑的详细解剖学的几何形状更无变化。我稍后在回应某些持怀疑态度的评论时，再返回到受试者之间的这种不变性。

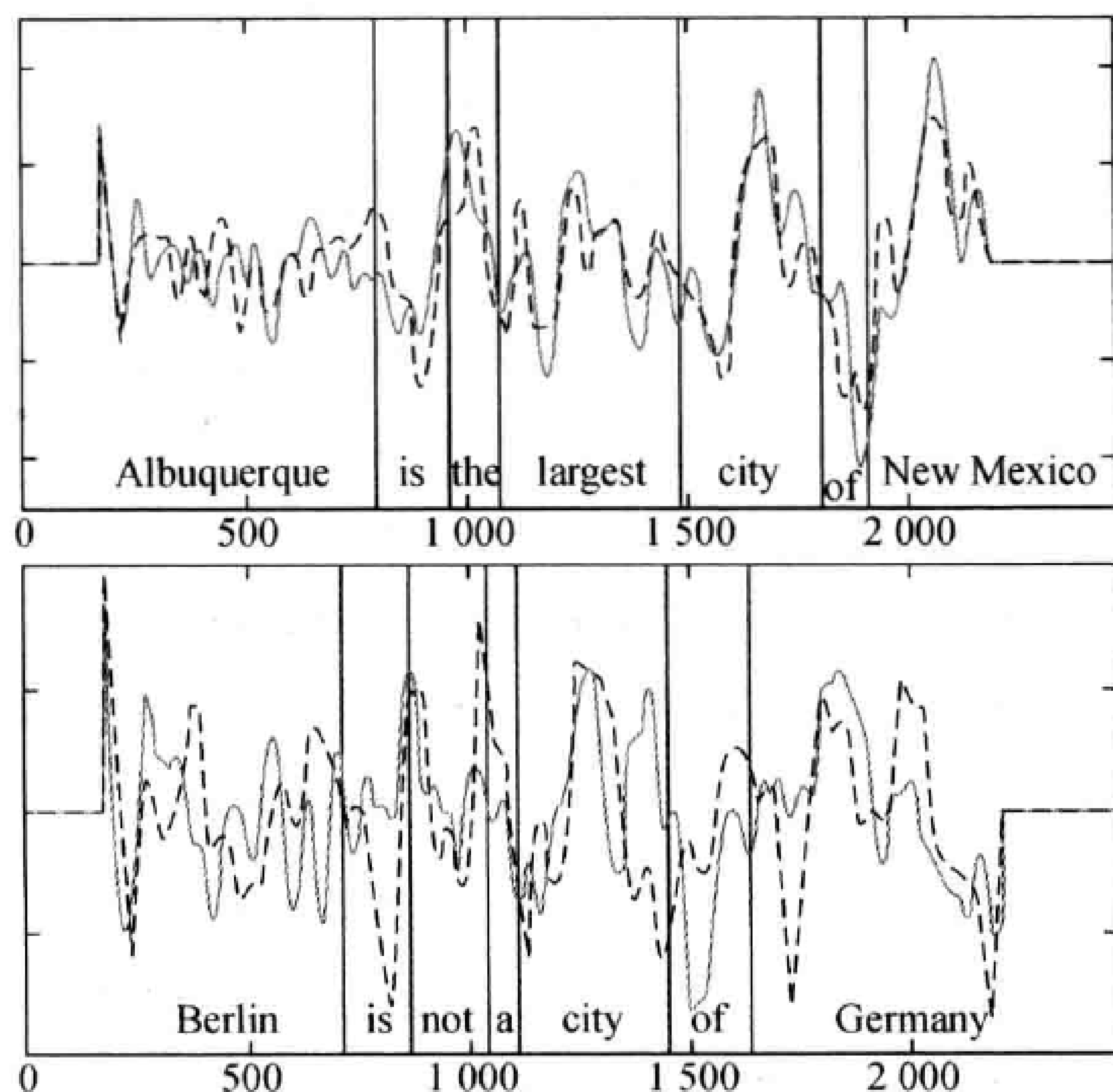


图 8　原型(灰线)和由受试者 S32 能正确分类的拟合最好的句子(上面的图形)和拟合最差的句子(下面的图形)产生的测试样本(黑虚线)。横坐标上用毫秒表示句子开始后的时间测量。

一百个句子。我现在转向第二个更近的实验，在这个实验中，向受试者直观地呈现出一百个不同的地理句子(苏佩斯和翁等人提供)。我这里只关注一位受试者(S32)从一百个句子中正确地识别出九十三个句子的显著结果。运用上面描述的方法，单独一位受试者所达到的最好的识别率是 93%，即一百个测试样本中识别出九十三个。获得最好结果的参数为 $L = 1.25$ Hz，$W = 21.25$ Hz，$s = 180$ ms 是看见每个句子的第一　452

个词开始显示之后的时间，$e = 2\,200$ ms 标志着用最小二乘法拟合标准进行记录的结束时间。最好的双极传感器是 C4 - T6。在图 8 中，我们表明了以最小二乘法标准为测度，正确识别出九十三个句子时得到上图的最好拟合和下图的最差拟合。最差拟合的平方和是最好拟合的平方和的三倍多。

视觉图像及其名称之间的不变性。第三个实验表明，在一台计算机屏幕上产生的熟悉形状的视觉图像(比如，一个圆或三角形)，非常类似于由相应的词产生的大脑图像(Suppes, Han, Epelboim and Lu 1999b)。这种非常令人惊奇的结果极大地强化了思想中如何有一般理念的经典解答。下面是在 18 世纪的哲学史中，贝克莱和休谟强烈批判洛克的抽象概念或一般思想概念的一段著名插曲。贝克莱在《视觉新论》(1709/1901)一书中说到了此事：

> 所有的一般真理都是关于普遍的抽象理念的，这确实是现代哲学家和古代哲学家的一个信条；我们被告知说，如果没有这个信条，根本不可能有科学，一般的几何命题都是无法证明的。但这是容易明白的，我的确认为，为了达到我当前的目标，有必要表明，几何中命题和证明可能是普遍的，尽管提出这些命题和证明的那些人从来没有想到抽象三角形或圆的一般理念。
>
> 我在经历了考虑理解三角形的一般理念的反复努力与阵痛之后，我发现，这种一般理念是完全不可理解的。确实，如果有人能使那种理念进入我的思想，那么，他一定是《人类理解论》的作者：他，因其论述的明晰性和重要性，迄今为止把他自己与大多数作者区别开来。因此，让我们看一下，这位著名作者如何描述一个三角形的一般理念或同

> 样的东西即抽象理念。他说,"它一定""既不是直角三角形,也不是长方形,既不是等边三角形、等腰三角形,也不是斜三角形;但它们全是同时又全不是。实际上,它连一点缺陷都没有;一种理念把几种不同的和不一致的理念的某些部分整合起来。"(*Essay on Human Understanding*, B. iv. ch. 7. s. 9.)他认为,扩展知识所需要的正是这种理念,它是数学证明的主题,没有它,我们无法知道关于三角形的一般命题。我能肯定,如果情况如此,对我来说,甚至了解最基本的几何要素也是不可能的:因为我没有能力表达我思想中像这里描述的那样一种理念。
>
> (Berkeley: pp. 188 - 189)

休谟在对贝克莱的思想作出英明的阐述和推广时,在《人性论》的前几页,标题为"抽象理念"的第Ⅶ节的第一段漂亮地表达了这个问题:

> **抽象的或一般的**理念提出了一个非常重要的问题,**不管它们在其思想概念中是一般的,还是特殊的**。一位伟大的哲学家对公认的这种特殊看法提出质疑,并断言,所有的一般理念只能是特殊的,附属于一个特定的术语,这个术语赋予这些一般理念更广泛的意义,并使它们回想起类似于它们的其他个别场合。当我把这一点看成是近年来文坛上作出的最伟大的和最有价值的发现之一时,我这里将努力通过某些论证来确证它,我希望这会使它确信无疑。 453
>
> (Hume, *Treatise*: p. 17)

尽管贝克莱和休谟没有讨论,我们也能证明,简单的色块就是这样情况。换言之,一块红色和"红"这个词在大脑皮层的听觉区产生了相似的大脑图像。

特别有意义的结果是这样的。通过对所有的受试者和试验样本求平均,我们用由简单视觉图像构成的刺激诱发的脑电图产生原型,用由确定视觉图像的听到或看到的词诱发的脑电图产生测试样本。我们对测试样本的脑电图的正确识别率是从60%到75%。我们的一般结论是,简单形状和简单色块产生的脑电图,与它们的口头名称所产生的那些脑电图惊人地相似。这种结论,与听觉记忆和视觉记忆的广泛的心理学研究一起,支持了贝克莱和休谟所猜想的解答。大脑,或者,如果你喜欢的话,是心灵,把(例如)三角形的单个视觉图像与*三角形*这个词联想在一起。正是这样一种联想网络,有可能在程序上替代洛克引入抽象理念的错误尝试。

由视觉图像产生的平均的和经过过滤的脑电图与说出该图像的名称所产生的平均的和过滤后的脑电图的比较,如图9所示。刺激(视觉图像或词)开始后的时间在横坐标上用毫秒表示。上面的图中的实的曲线是由在计算机屏幕上以蓝色背景的空白处显示的蓝色产生的脑电图原型。虚的曲线是由说出*蓝色*这个词所产生的脑电图测试样本。下面的图是由在计算机屏幕上显示的三角形产生的脑电图原型(实线)和由说出*三角形*这个词产生的脑电图测试样本(虚线)。两种情况都不完全匹配,即使重复同样的刺激,过滤后的脑电图也不会正好相匹配,因为大脑的电活动从一个时刻到另一个时刻以很多方式不断地发生变化。但尽管如此,存在着许多人类交流需要的不变性,甚至在这个早期阶段,我们就能识别出其中的一部分。

对结果的批评与回应。我首先概述这些批评的一般性。在许多脑成像实验中,数据是非常丰富而复杂的。因此,也可以用一种复杂程序寻找给定条件下的一个最优值。这个最优值(即这里的最正确的识别率)的搜索,类似于计算特定概率分布的一

个极值统计。这种类比的基础是,这种搜索可能相当于对一个零假设实验的许多次重复。这些重复要求计算适当的极值统计。此外,在寻找这样一个最优值时,有几个参数是被估算的,那么所找到的值的意义可能会受到挑战。这样一种挑战的基础是主张:对于大量数据和几个被估计的参数来说,在这个实验中,即使随机分配的有意义的标签,经过足够多次的随机重复,也仍然可能产生一种很好的预测结果。

454

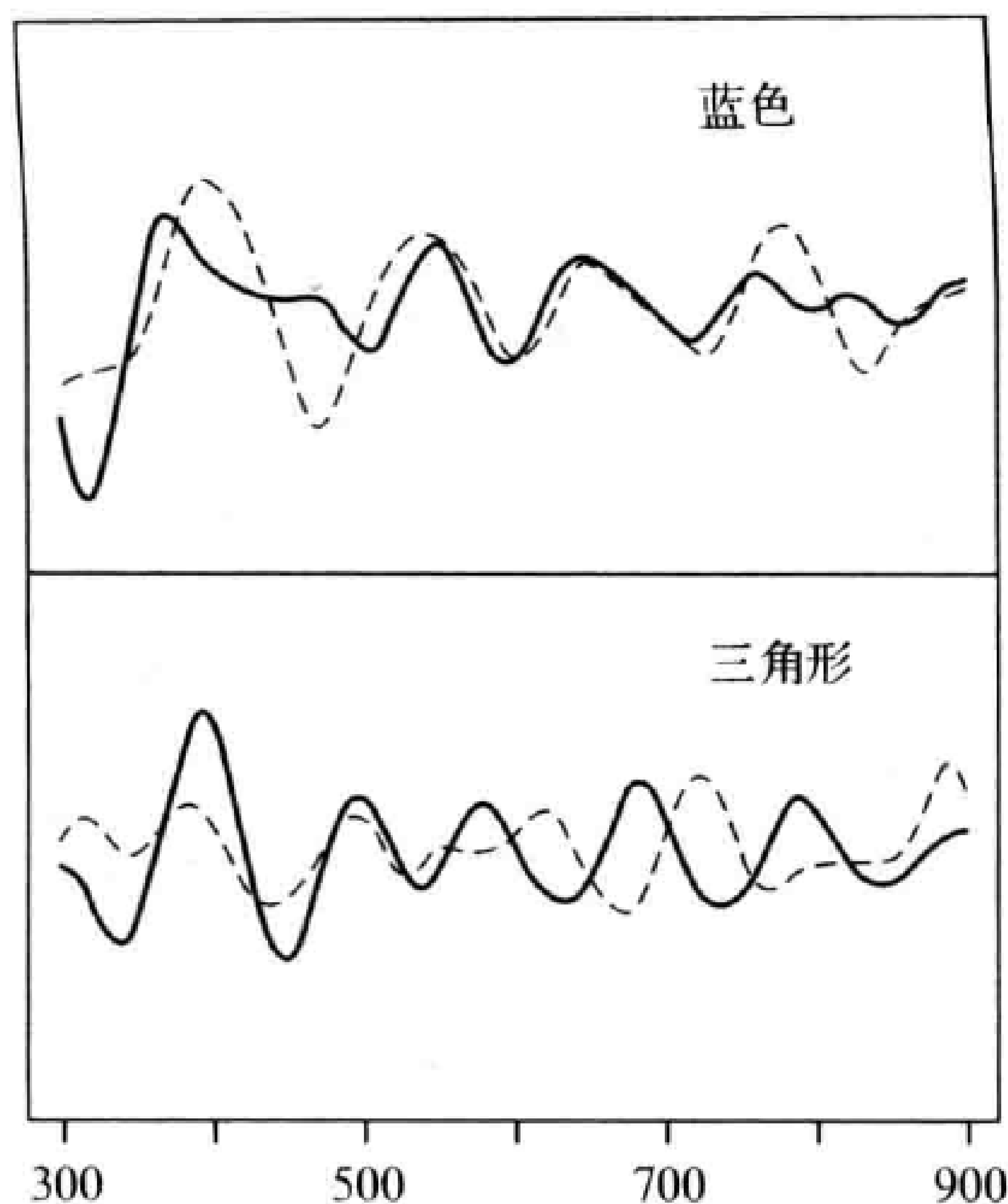


图 9　由视觉图像产生的滤波后的脑电波(实线)与说出该图像名称产生的脑电波(虚线)的比较

极值统计。我们用两种方式来应对这种批评。第一种方式是,在由标签的这些随机分配产生的二项式分布的零假设条件下,推导出极值统计分布。我们计算在科学上有趣的预测结果离极值统计分布的平均值的标准偏差有多少。物理学家通常把至少三倍或四倍的标准偏差接受为一个有意义的结果。其他科

学家和统计学家通常更喜欢用 P 值表示在零假设条件下观察结果的概率。这里我们报告了两种测度。

第二条进路意味着回答这些人的疑问:他们怀疑,即使在标签的一次随机排列之后,只含有一个参数(即随机地作出一种正确分类的概率)的零假设,也能适当地描述这些数据的结构。为了回应这些可能的怀疑,我们也计算对标签的五十次随机排列的采样的识别率。然后,我们用贝塔分布(beta distribution)来拟合给定实验条件的每个这样的采样结果,与由零假设产生的相应的极值统计分布进行比较。

455

我们首先推导零假设条件下的极值统计。

设 $p=$ 一次成功的概率,即在我们的试验中正确分类一个样本的概率,以及 $q=1-p$。设 X 是随机变量,它的值是在 n 个独立试验中成功的次数 k。至少 k 次成功的概率是

$$P(X \geqslant k) = \sum_{j=k}^{n} P(X=j) = \sum_{j=k}^{n} \binom{n}{j} p^{j} q^{n-j} \tag{3}$$

现在,我们重复由二项式分布决定的实验。因此,我们对 n 次独立试验进行 r 次独立重复。对于 r 次重复来说(在一百个句子的实验中 $r=21\,000$),表示极值统计的随机变量是

$$Y = \max(X_1, \cdots, X_r) \tag{4}$$

设 $P(Y \geqslant k)$ 是 Y 在 r 次重复中至少有一次大于等于 k 的概率,即我们感兴趣的极值统计。那么显然,

$$\begin{aligned} P(Y \geqslant k) &= 1 - P(X < k)^{r} \\ &= 1 - \left[\sum_{j=0}^{k-1} \binom{n}{j} p^{j} q^{n-j}\right]^{r} \end{aligned} \tag{5}$$

我们也需要 Y 在理论上的密度分布,与后面的各种经验结果作比较。这不难从(3)计算出来。

$$P(Y = k) = P(Y \geqslant k) - P(Y \geqslant k+1)$$
$$= P(X < k+1)^r - P(X < k)^r$$
$$= \left[\sum_{j=1}^{k}\binom{n}{j}p^j q^{n-j}\right]^r - \left[\sum_{j=1}^{k-1}\binom{n}{j}p^j q^{n-j}\right]^r \quad (6)$$

从(6)[①]我们能算出极值统计 Y 的平均和标准偏差。

第二,我们报告了用在(0, 1)上的贝塔分布拟合极值统计的经验样本的结果。贝塔分布的密度 $f(x)$是

$$f(x) = \begin{cases} \dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}, & a,\ b > 0,\ 0 < x < 1 \\ 0 & \text{其他情况} \end{cases} \quad (7)$$

这里,$\Gamma(a)$是伽马函数(gamma function)。如果 Z 是一个具有贝塔分布的随机变量,那么作为参数 a 和 b 的简单函数给出它的平均值和方差。

$$\mu_Z = E(Z) = \frac{a}{a+b} \quad (8)$$

$$\sigma_Z^2 = \text{Var}(Z) = \frac{ab}{(a+b)^2(a+b+1)} \quad (9)$$

随机变量 Z 具有等于或大于$\frac{k}{n}$的值的概率是

$$P(Z \geqslant \frac{k}{n}) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_{\frac{k}{n}}^{1} x^{a-1}(1-x)^{b-1}dx \quad (10)$$

计算 $P\left(Z \geqslant \frac{k}{n}\right)$,对这种分布的极值尾部来说,是困难的。在

① 原文为(4),这是王瑞博士在译校本小节时,发现原文有误,并改正为(6)。——译者

456 某些情况下，我们用了不是最可能的但容易计算的数学上严格的上界，也就是，恰好是高度为 $f\left(\frac{k}{n}\right)$ 的包含了 $f\left(\frac{k}{n}\right)$ 的右边分布尾部的长方形的面积：

$$P(Z \geqslant \frac{k}{n}) \leqslant f\left(\frac{k}{n}\right)(1-\frac{k}{n})$$

这里 $f\left(\frac{k}{n}\right)$ 是通过(5)来定义的。

计算极值统计。我们从用一百个句子的第二个实验开始。作为对零假设的一种检验，我们通过抽样五十个随机排列，构造极值统计的一个经验分布。应注意的几点是：

1. 一百个语句“标签”的一种排列是从总体上一百种可能排列中随机地抽取出来的，并用这种排列重新标记句子的测试样本。

2. 像在处理正确标记受试者 S32 的数据时那样，现在完全是为每对双极传感器搜索相同的参数网格(L, W, s, e)，通过傅里叶分析、滤波和选择时间区间(s, e)，来获得随机标签分配的最佳分类或识别率。对于一百个句子的实验来说，每个随机排列的测试网格上的节点数是，$L \times W$ 为 $7 \times 10 = 70$，$s \times e$ 为 $5 \times 4 = 20$ 以及传感器个数为 15，因此，从零假设的立场来看，重复次数是 $70 \times 20 \times 15 = 21\,000$。

3. 不断重复标签排列的这种随机抽样和计算识别结果，直到获得五十种排列的样本为止。

在图 10 中，我表明了对受试者 S32 的数据的五十种标签排列的样本的平均值 m 和标准偏差 s 的累计计算。对于五十个所有样本来说，平均值 $m = 6.04$，标准偏差 $s = 0.77$。在图 11 中，我表明了：(i)具有 $n = 100$，$p = 0.01$ 和 $r = 21.000$ 的零假

设极值统计 Y 的频率分布,(ii)从对 $r = 21.000$ 的五十次采样获得的最大成功数的实验统计柱状图,以及(iii)拟合的贝塔分布。从图 11 明显地看出,对于 S32 来说,一百个句子中有八十

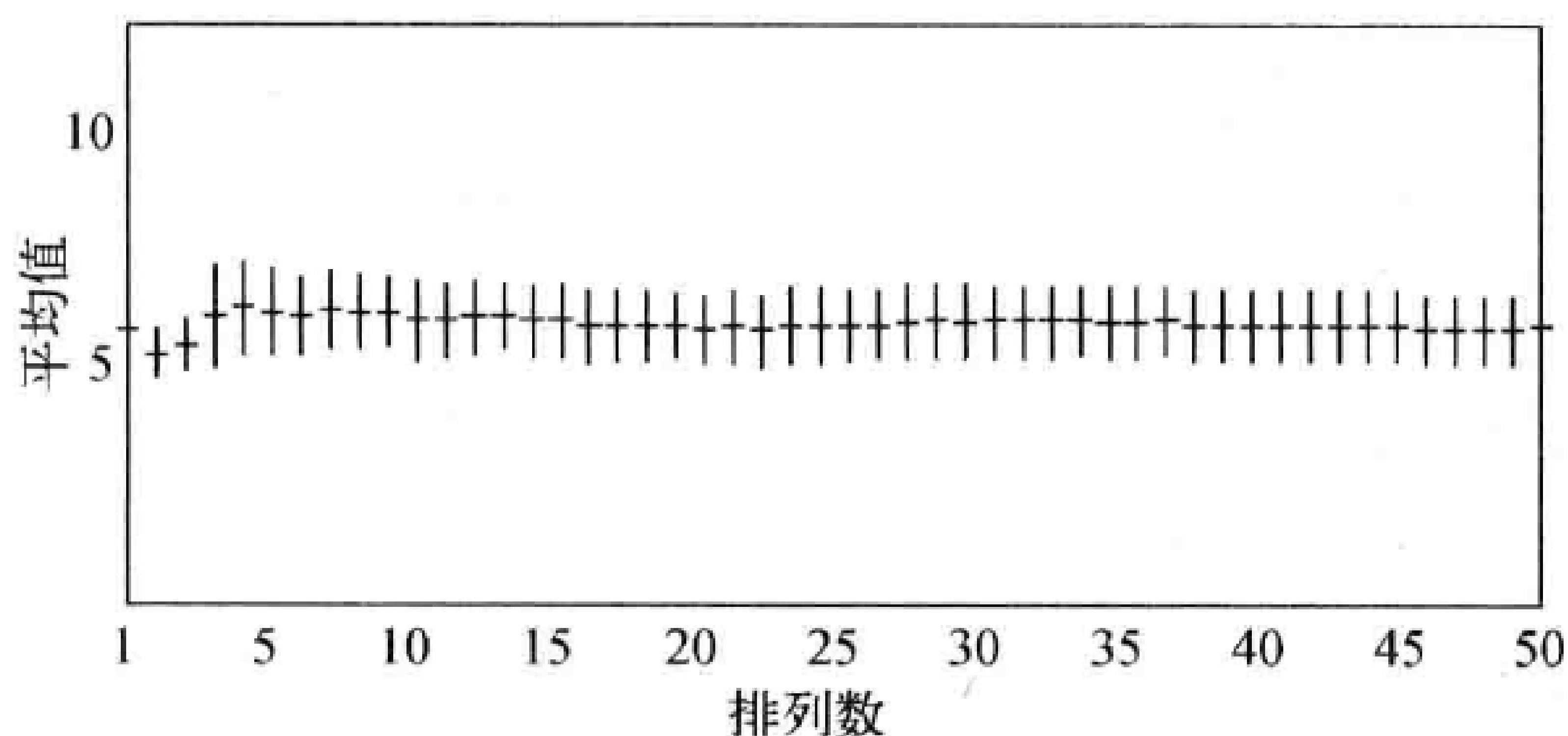


图 10　随机排列样本的识别率的累计标准偏差和平均值

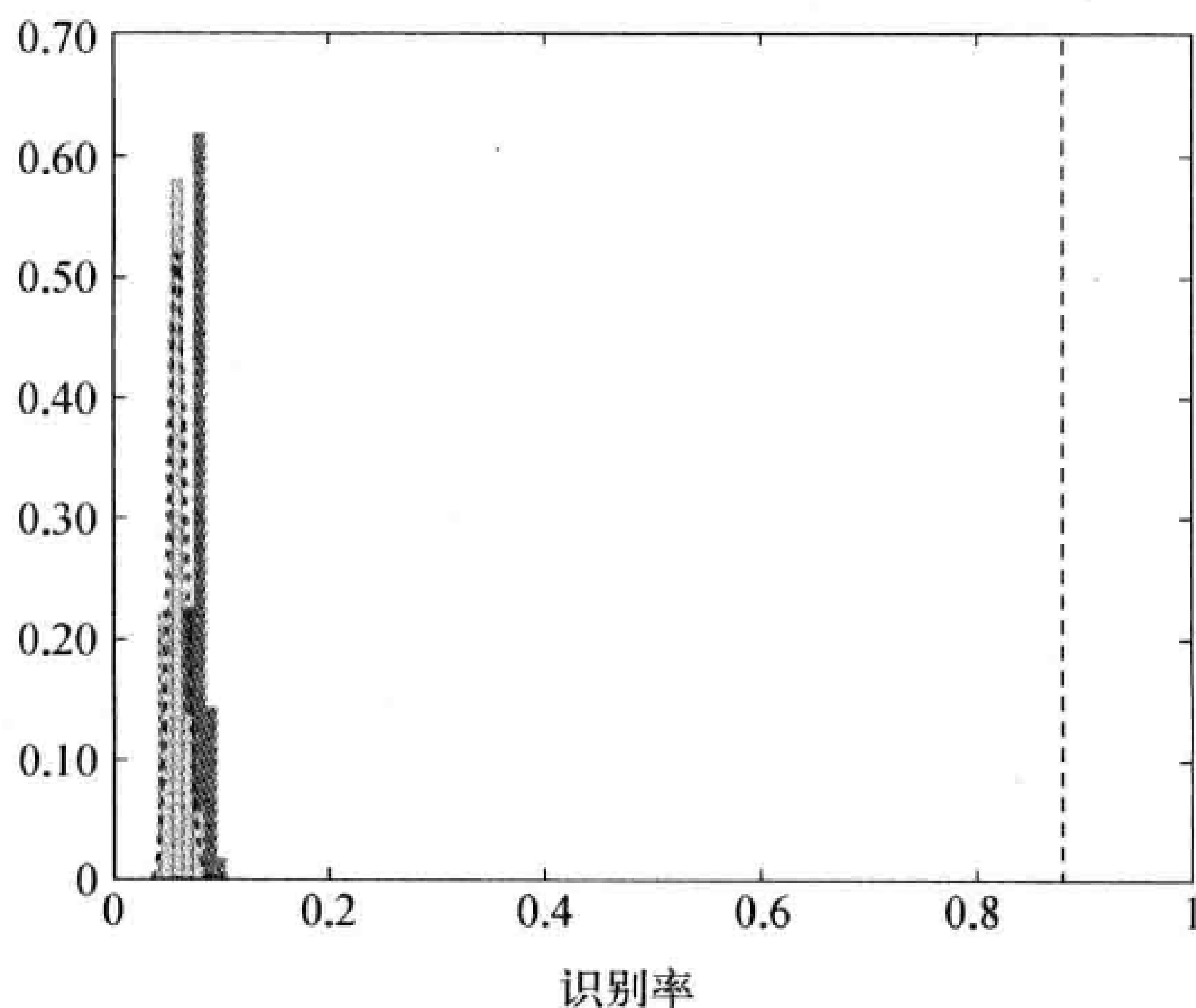


图 11　零假设极值统计 Y 的频率分布(黑色区域),五十种随机排列样本的柱状统计图(浅色阴影区),拟合的贝塔分布(虚线),以及右边(垂直短画线)S32 的识别率。

457 多个句子的正确分类要么与极值统计 Y 的分布不一致，要么与基于这些标签的五十种随机排列的样本网格计算所估计的贝塔分布不一致。贝塔分布比 Y 分布稍微更拟合得好些，这一事实是不足为奇的，因为后者没有估计自由参数。下一段讨论有更大 r 的一种更精细的搜索，它产生了一百个句子中正确分类出九十三个句子的更好结果。

也许令人惊奇的是，零假设极值统计 Y 的平均值 $\mu = 6.95$ 稍微大于实验样本分布的平均值 $m = 6.04$。有三点是值得注意的。首先，实验样本的标准偏差 $s = 0.77$ 大于极值统计 Y 的标准偏差 $s = 0.67$。我下面评论这种差别。其次，我在图 12 中表明，当 r 增加一个或几个数量级时，对于 $n = 100$, $p = 0.01$ 以及后面用到的 p 和 n① 的其他值，零假设极值统计 Y 识别率的增长率。正如所能看到的那样，在零假设条件下，正确识别的增

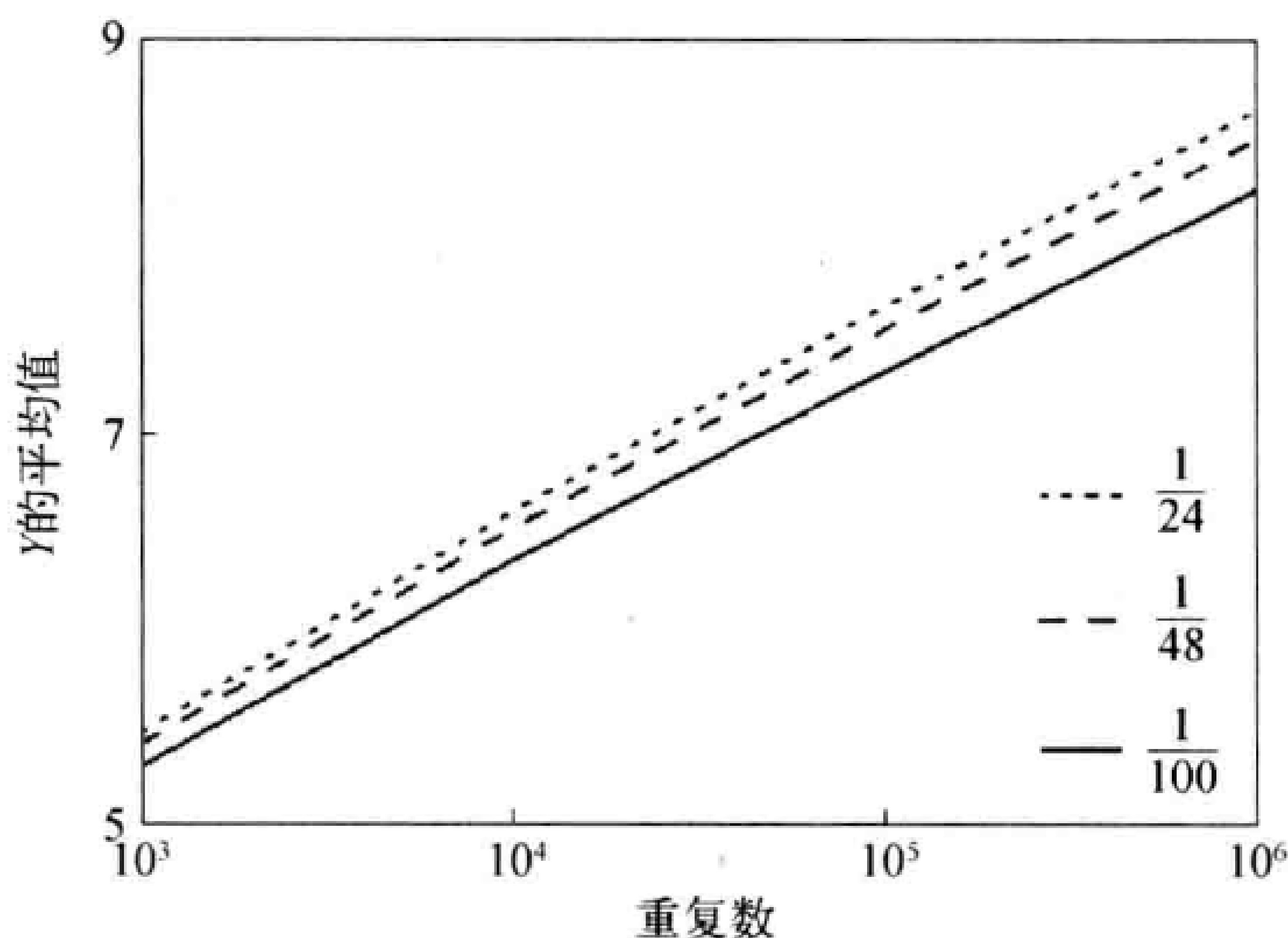


图 12　作为重复次数 r 的一个函数的零假设极值统计 Y 的平均值

① 原文为 r 有误，这是作者的学生王瑞博士在校译时指出的，并做了相应的改正。——译者

长率是缓慢的。作为一个重要的例子,我们通过扩展搜索 S32 的数据网格进行改善。我们不用完整的网格,而只是在一个有前途的方向上进行改善和扩展。在所有方向上进行广泛比较,r 的数量极是 10^7,即重复 10 000 000 次。因此,我们为 r 的这个大值(它比我们实际计算出的任何一个值都大得多)计算 Y 的零假设分布。即使对于这个大的网格来说,当 $r = 10^7$ 时,极值统计 Y 的平均值只到 $\mu = 9.60$。由于标准偏差现在减少到 0.62,所以,93 和 9.60 之间距离的标准偏差单元的数是 134.5, 458 大于以前的数。①

第三点聚集我们为了拒绝零假设报告的显著性水平或 p 值。$k = 93$ 的结果(即是离零假设极值统计 Y 的平均值 134.5 倍的标准偏差)的 p 值非常小,至少是 $p < 10^{-100}$。这个实验的其他可能方面必须是完美的,才能支持这样一个 p 值。(我们确实在运行不同程序的两台不同计算机上检查了 93 的计算。)因此在这里和后面的其他情况下,除了报告实际离平均值的标准偏差数之外,对于这些很小的值来说,我们只报告不等式 $p < 10^{-10}$。

我们也检查了用不等式(11)②的严格上界为拟合的贝塔分布所计算的 $P(X \geqslant 93)$,也在 $p < 10^{-100}$ 的阶数之上。这进一步支持了这种观点:后面用到的 p 值的不等式,也就是 $p < 10^{-10}$,是很保守的。

---

① 在反思这些结果时,重要的是记住,物理学家通常很高兴在经典或零假设的预言与新观察到的结果之间有 6 或 7 倍的标准偏差的分离,例如,在 7.2 节讨论的那类量子纠缠实验中。在我们的大脑实验中所获得的显著性水平,以及在从预计的零假设结果到观察到的结果的标准偏差中的很大距离,在心理学或神经科学实验中,至今还很少见。

② 原文是(9),作者的学生王瑞博士在校译时发现原文是错误的,并加以更正。——译者

对早期研究的分析。我已经强调了对所有的受试者和试验求平均,可以改善预测结果。在第二个实验中,一位受试者的93%的最好结果,促使我们重新考虑早期实验中最好的个别结果。在我们完成的每个实验中,至少对一位受试者的脑电图的分析产生了大于90%的一种正确分类,已经提及的四十八个句子的实验例外,其结果是77%[Suppes, Han, Epelboim and Lu (1999a),这个77%是作为79%报告的,因为用了更精细的网格]。表6总结了不同实验中的最佳受试者的结果。

459

除了一个例外,表6所示的 p 值都是很显著的,从实验工作的大多数标准来看,更是如此。这个例外是视觉图像实验,在这个实验中,八个简单的视觉图像作为刺激显示出来。如表6所示,对于四位受试者来说,我们能够对所有的八个脑电图进行正确分类,但对于零假设来说,100%的完美结果只在 $p < 0.01$ 的水平上是显著的,因为经过足够的重复之后,由于 $\mu = 6.28$,所以,最佳猜测在零假设条件下也非常好。

后面这种观点的实验设计的教训是显而易见的。像在所描述的大脑实验中那样,如果数据众多而复杂,还要求广泛地搜索最优参数,那么,在零假设条件下,一个正确反应的概率 p 应该很小。图12用图形说明了这一点。当 p 小时,重复次数可以很大,但不对 r 次重复的极值统计 Y 的平均值造成很大的影响。也正如从表6中能够看到的那样,当一百个句子实验的二项式参数 $p = 0.01$ 时,即使在一百个试验的零假设条件下,重复一千万次,也只能把 $E(Y)$ 稍微增加到9.60。为了使这个论点很引人注目,我们说,当每个试验的重复率是一秒时,经过足够次的重复获得 $E(Y) \geqslant 93$ 所花费的时间,比现在所估计的宇宙年龄还大。

我在前一段指出,p 应该很小,但这太简单。在表6的前两

表 6　例外的识别率[a]

实　验	受试者	成功的次数	机会概率	正确的百分比	重复数 r	统计值 Y			显　著　性			
						μ	m	σ	s	$\#\sigma$	$\#s$	p 值
7 个视觉词[1]	S1	32 of 35	1 of 7	91	2 925	13.40		0.92		20.2		$<10^{-10}$
7 个听觉词[1]	S3	34 of 35	1 of 7	97	3 600	13.55		0.91		22.5		$<10^{-10}$
12 个句子[2]	S8	56 of 60	1 of 12	93	60 480	16.10		0.90		44.3		$<10^{-10}$
24 个视觉句子[3]	S18	24 of 24	1 of 24	100	30 800	6.83	5.64	0.63	0.79	27.3	23.2	$<10^{-10}$
48 个视觉句子[3]	S26	38 of 48	1 of 48	79	30 800	7.04	6.20	0.63	0.94	49.1	33.8	$<10^{-10}$
8 个视觉图像[4]	4Ss	8 of 8	1 of 8	100	95 550	6.28	4.92	0.46	0.69	3.7	4.5	<0.01
100 个视觉句子	S32	88 of 100	1 of 100	88	21 000	6.95	6.04	0.67	0.77	121.0	106.4	$<10^{-10}$
100 个视觉句子	S32	93 of 100	1 of 100	93	10^7	9.60		0.62		134.5		$<10^{-10}$

1. Suppes, Lu, and Han (1997); 2. Suppes, Han and Lu (1998)
3. Suppes, Han, Epelboim and Lu (1999a); 4. Suppes, Han, Epelboim and Lu (1999b)

a 第一列是实验清单,最后两个条目是一百个句子的实验。第二列是受试者名单,根据苏佩斯、卢和哈恩的文章(Suppes, Lu and Han, 1997)中第一次报告的实验连续编号。第三列表明,从所有呈现的样本中获得成功识别的最大测试样本的个数。第四列是一个正确分类的机会概率,它只是 1 除以原型的个数。第五列记录的是从第三列计算出的正确的百分比。第六列表明了在计算极值统计的特殊实验中所用重复次数 r。这个数 r 也是在正确标记的初始搜索的网格中用到的重复数。第七列记录了零假设极值统计 Y 的平均值 μ。第八列记录了我们作出这种计算的实验的极值统计的实验样本的平均值。第九列表明了极值统计 Y 的标准偏差 σ。第十列是实验样本的相应的标准偏差 s。第十一列记录了数 $\frac{k-\mu}{\sigma}$,即在第三列中记录的成功数 k 离极值统计 Y 的零假设分布的平均值 μ 的标准偏差数。第十二列是相应的实验样本数。第十三列表明了在等式(6)给出的极值统计 Y 的分布下,所观察到的成功数 k 的 p 值的一个保守的边界。在视觉图像实验的四个受试者的情况下,m 和 s 是四者的平均。描述每个实验时的上角标是已发表研究成果的参考文献。[除了在 Suppes, Han, Epelboim and lu 文章中(1999b)的四位受试者之外,每个受试者的 EEG 传感器或对偶双极传感器以及最优滤波器如下:

S1: T6, 1－10 Hz; S3: T3, 3－11 Hz; S8: C4－T6, 2.5－9 Hz; S18: P4－T6, 0.5－10 Hz; S26: C4－C6, 1－15 Hz; S32: C4－T6, 1.25－22.5 Hz。最优参数通常并不惟一。]

个实验条件(即七个视觉词和七个听觉词)中用到的另一种解决办法是增加测试样本的个数。在这两种条件下，$p=\frac{1}{7}$，但测试样本的个数是35，正如从这个表中能看出的那样，在比10^{-10}更好的水平上拒绝了零假设。使用类似方法，通过把测试样本的个数从8增加到24，重新分析$p=\frac{1}{8}$的视觉图像实验的数据，产生了拒绝零假设的某些更好的水平。细节后面报告。

更可疑的问题。像在所有的详细实验领域内一样，不可能存在一个明确的时刻，在此之后，由于已经解决了所有的相关问题，所以，不需要进行进一步的实验。伽里森(Galison)对物理学中几个重要的实验研究程序进行了详细的研究。他分析的主要方面似乎有可能应用于许多其他科学领域。他说，

> 在任何实验规划的不同检验和论证中，实验者必须潜在地或显在地决定，他们的结论在其他条件不变时，即在所有其他因素都是相等的时候，是站得住脚的。而且，他们必须这样做，尽管事实是，结束一种实验证实不是、也不可能是，完全基于一个闭合的程序集。……仪器的某些操作、计算、假设和论证为实验者提供了信心：它们是什么呢？……实验者什么时候敢承认一种结果是真实的呢？他们什么时候断言，计数器的脉冲或一个图形中的波峰超越了仪器或环境的假象呢？简而言之，实验是如何结束的呢？
>
> (Galison：pp. 3 - 4)

在当前的大脑实验语境中,问题实际上不是实验什么时候结束呢?而是对实验数据的计算什么时候完成呢?我考察了更加不同的但典型的三种质疑,这三个问题是针对新的研究提出的,这种新研究的有效性得到了强有力的统计支持。 461

在含有四十八个句子的第一个实验中的其他对。有些怀疑者对我们评论说,我们只是幸运地对我们分析的受试者进行了特殊的2元划分。因此,我们根据与表6一样的最优值,对九位受试者的所有五百一十次的2元划分进行了测试,而没有尝试搜索网格上所有的节点(Suppes, Wong et al. 即将发表)。在图13中,我们显示了五百一十次划分结果的柱形统计图。对于在五百一十种可能性中除四种可能性之外的所有可能性来说,这些结果的显著性水平是 $p < 10^{-10}$,而且,这四种可能性之一具有 $p < 10^{-7}$ 的显著性。最好的结果是从四十八个中识别出四十六个,这对几种划分都成立。因此,当前更彻底的统计分析

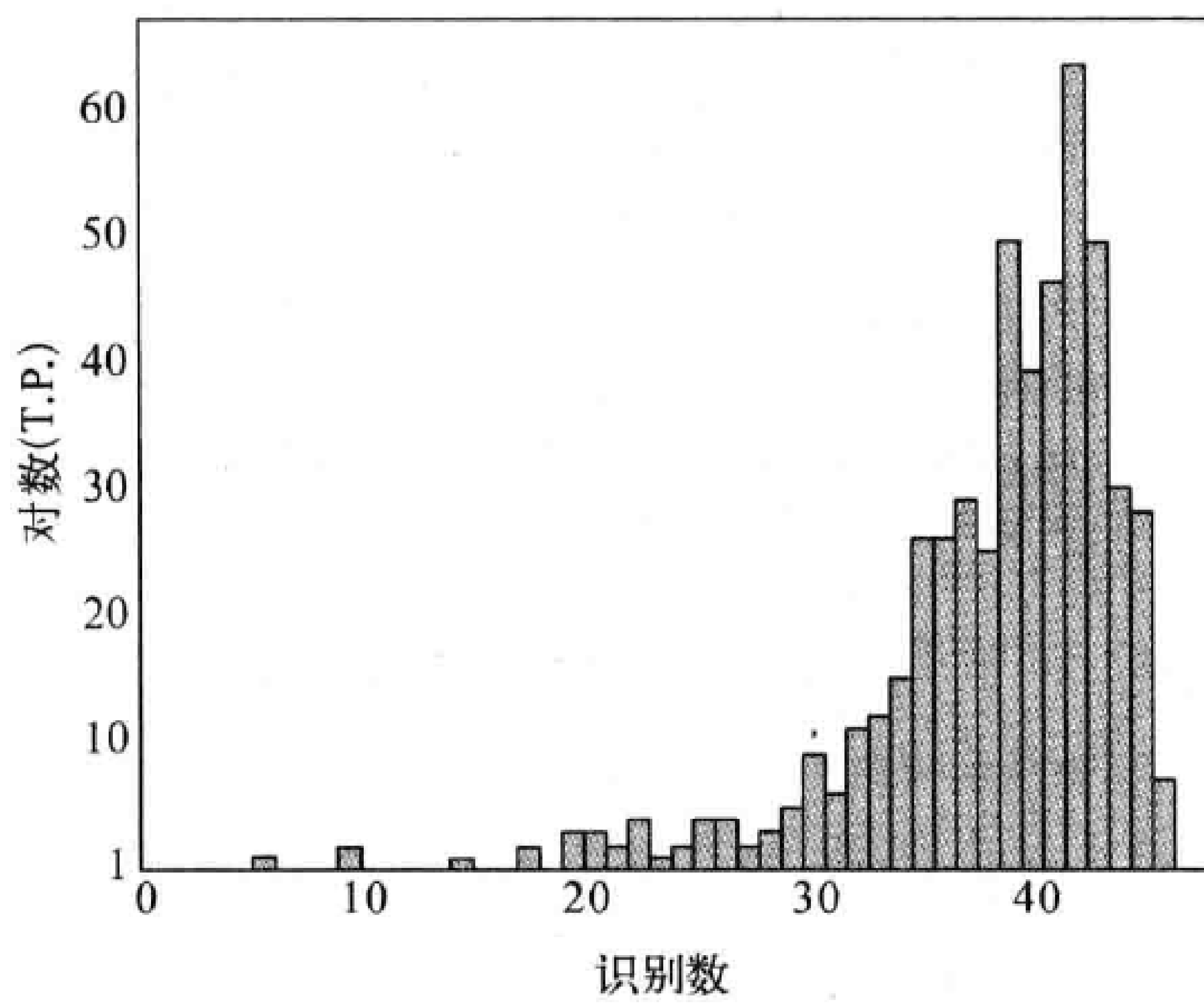


图13 在四十八个句子的实验中,对九位受试者的五百一十次2元划分的正确识别率的柱形统计图

更坚定地支持了前面的研究结论:受试者之间脑电图的不变性。图 14 表明了相同数据的另一种看法,在横坐标上标出每个 2 元划分的原型中的受试者的人数,纵坐标上标出对每类原型的四十八个句子作出正确分类的平均数。令人吃惊的是,当原型只有一位受试者或除一位受试者之外所有的其他受试者(即八位受试者)时,平均结果很好($p < 10^{-10}$)。这个证据相当有说服力,我们最初选择的划分不只是某种幸运。

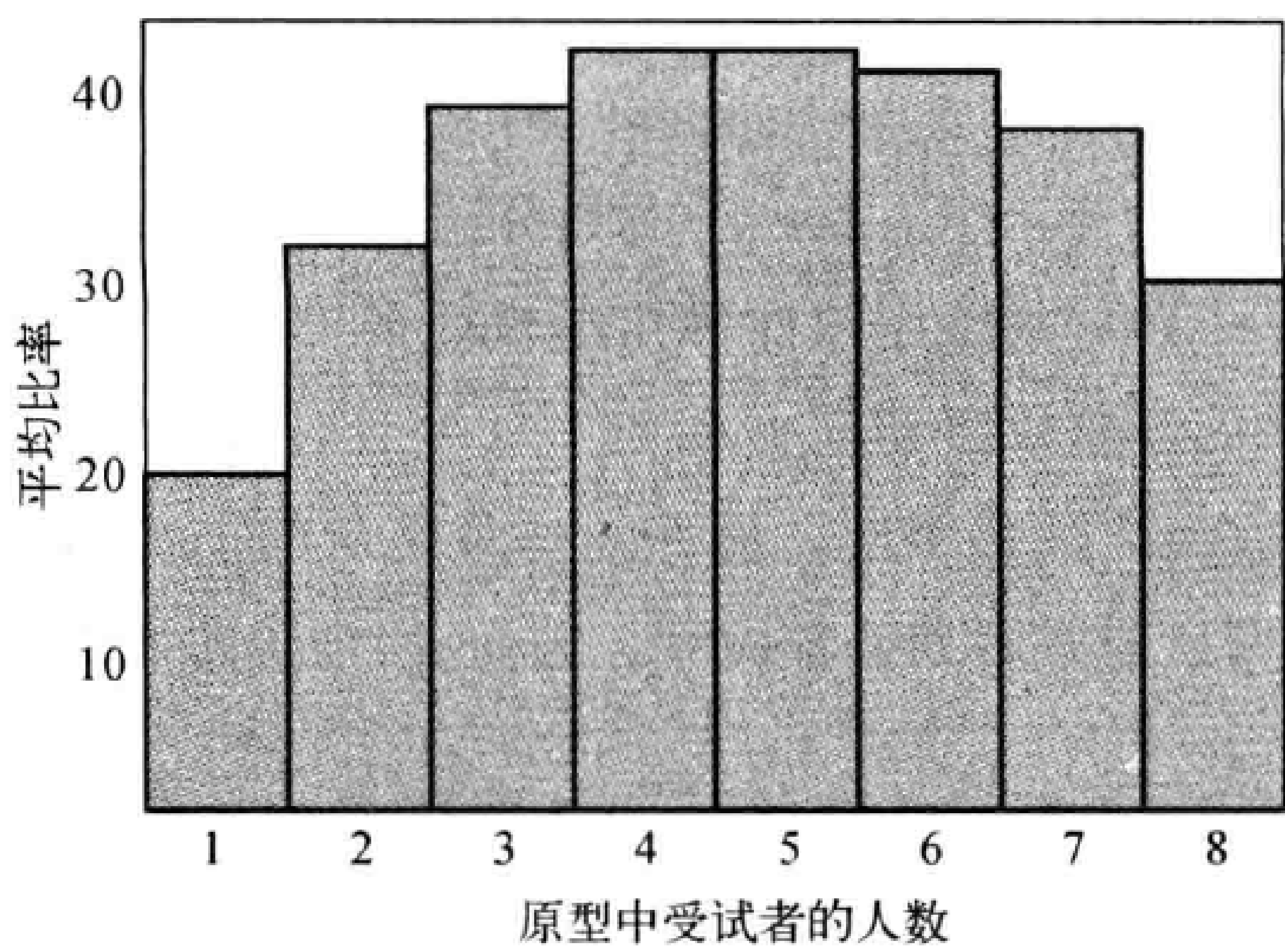


图 14　对五百一十次 2 元划分的正确识别率的平均数的柱形图,标记 1-8 代表原型中受试者的个数。因此,比如,横坐标上的 1 对应于原型中只用了一位受试者的所有 2 元划分。

对含有一百个句子实验的计时假设的检验。上面的第二个实验报告中得到了 93%的高识别率,表 6 总结了前面的结果,在与同事讨论这个结果时,有几个人怀疑地建议说,也许我们的识别率恰好来自对词在不同句子中的视觉呈现的不同计时。这些句子每次只在计算机屏幕的中心出现一个词,同时,每个视觉词出现的开始时间与相应句子的听觉呈现的开始时间一样。在

462

视觉上每次呈现一个词避免会干扰脑电图的许多令人棘手的眼睛转动，这种方法也被看成是读出详细内容的一种有效的捷径(Rubin and Turano, 1992)。一个句子的每个视觉词的持续时间也与在几毫秒内听到这个词的持续时间相匹配。

大多数视觉词一开始诱发反应电位的出现，支持了这种计时想法。为了检验这种想法，我们所用的识别模型是：它只取决于一个句子中的每个词的脑电图反应的起始段(Suppes, Wong et al. 即将发表)。这种模型用两个不同参数取代了表示时间区间的两个参数 s 和 e。第一参数是 α，它是在每个句子中每个词一开始与相对应的大脑皮层中的脑电图的开始之间估计的时间滞后。第二个参数是 β，它是每个词从一开始显示其词长的比例，表示在信号到达大脑皮层的延迟时间 α 之后，在原型中用来识别的部分。因为词长和句长是可变的，所以，我们通过除以所用的观察采样个数来归一化最小平方的计算。如果在识别过程中，只有计时，而不是整个词的表征，是重要的，那么只需要一个词的起始段的一小部分，基本上是包含词的开端诱发的反应电位的起始段。另一方面，如果在成功的识别中使用了整个词的表征，根据我们的最小平方标准，那么 β 越大，识别越好。为了把 β 调整到所显示的每个词的时长，我们把 β 表示为每个句子的词 i 显示的时长的十进倍数。最好的预期结果是，当 $\alpha=200$ 毫秒和 $\beta=1.25$ 时，识别率为 92%。作为 $0.125\leqslant\beta\leqslant2.00$ 的一个函数的识别率如图 15 所示。正确的识别率随 β 单调增加直到 $\beta=1.25$，然后，在 $\beta=1.50$ 之后缓慢减小。这些结果支持了两个结论。第一，计时是重要的。对于 $\beta=0.125$ 的 45%的识别率远远大于一种偶然结果。但是，第二，脑电图对一个句子中的词的表征越完整，越有利于达到识别目标。 463

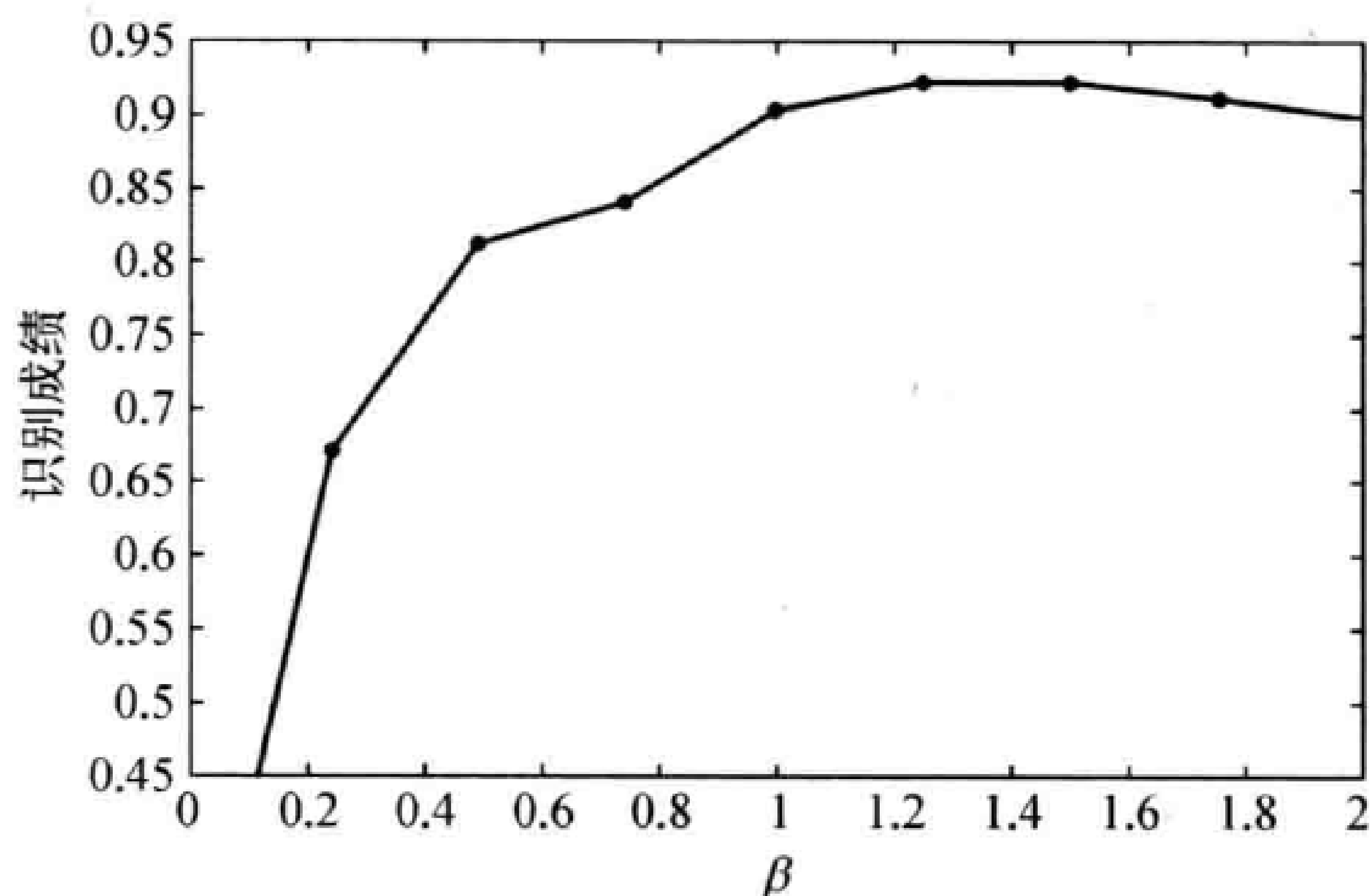


图 15　运用在一个句子中每个词开始后脑电图数据的不同起始段的分类结果。在一个句子中词 i 一开始 α 毫秒后，片段 i 开始，然后，在 $\beta \cdot l_i$ 毫秒后，片段 i 结束，这里，l_i 是词 i 的视觉呈现的长度，用毫秒表示。

在视觉图像实验中的截尾数据。继续计算下去的一种疑问是关于伪差问题。也许，引人注目的统计显著性水平归因于数据中的某些伪差。①现在，在 EEG 研究中，伪差问题有很长的历史。一个主要来源是眨眼和眼球跳动或眼动跟踪，另一个主要来源是周围环境的电流，主要由于在通常做实验的任何一个建筑物内标准交流电 60 Hz 的振荡，还有另外一个来源归于观察和记录脑电图的仪器。这个清单很不全面。几个不同的视角研究这个主题的文献有许多，但多得无法在这里概述。

464　我只限于描述，我们如何运用一种相当熟悉的统计方法，而

① 已知在第二个实验中，随机排列的极值统计的平均值接近于零假设的低平均值，这个实验的 93％的正确分类，或者，事实上，具有 $p < 10^{-10}$ 的其他任何实验结果，都不可能是由伪差造成的。但是，排除伪差仍然是一个重要话题，不是应对所知甚少的疑问，而是为了改进分类结果，下面的一个例子就是这种情况。

不是对在眨眼或其他伪差下记录的数据进行肉眼检查。[①]这种方法就是截尾数据，只引入一个自由参数使得截尾最佳——在已经描述的使正确识别率最大化的意义上的最佳。

设 X_{ik} = 在试验 k 中的第 i 次观察，设 ω 是产生一个原型或测试样本求平均计算所需要的试验次数。于是，

$$m_i = \overline{X}_i = \frac{1}{\omega}\sum_{k=1}^{\omega} X_{ik}$$

我们以同样的方式计算方差 s_i^2。设 α 是截尾的自由参数，使得如果

$$| X_{ik} - m_i | > \alpha s_i$$

从所平均的数据中排除在试验 k 中的第 i 次观察。计算任务是寻找使分类结果最优的 $\hat{\alpha}$。在这里举的这个例子中，我们运行 s_i 的 20 个值的一维网格，来近似计算最优值 $\hat{\alpha}$。

这个实验是已经描述的视觉图像实验，即与 18 世纪关于抽象理念的争论相关的实验，对这个实验来说，极值统计不太显著。正如我前面指出的，这些数据以一种直接的方式支持了贝克莱和休谟的怀疑论的观点，但是，统计支持并没有太大的说服力。因此，我们重新分析这些数据，创建二十四个而不是八个测试样本，我们还对只会说汉语(普通话)的四个人做实验，用汉语和英语中的听觉词或视觉词，确认语言部分。苏佩斯和翁等人(即将发表)的报告中包括了这个实验的细节。

表 7 表明了跨模式分类的显著性结果。前两个条件是针对最初用英语的实验的。用女声听觉刺激的脑电图(AWF)代表

---

① 在我们后来进行的四万次试验的更大实验中，试图用传统的观察方法检测伪差，是不切实际的。

原型,用视觉图像刺激的脑电图代表测试样本,截尾后,从二十四个测试样本中正确分类出十五个,改进了在不作截尾时从二十四个测试样本中正确分类出十一个的结果,所导致的显著性水平为 $p<0.016$。对调原型和测试样本,产生了更好的截尾结果,从二十四个测试样本中正确分类出十六个,显著性水平为 $p<0.001$。注意,这里的显著性水平是保守的,是基于全部网格搜索的,相当于 $r=1\,470\,000$,为极值统计零假设条件下的重复次数。

表 7　在截尾数据的视觉图像实验中的跨模式结果

实验	R	Y μ	σ	显著性	
				$\#\sigma$	p 值
英语					
AWF/VI	11	11.57	0.692	0.8	≈0.500
截尾	15	13.01	0.642	3.1	<0.016
VI/AWF	11	11.57	0.692	0.8	≈0.500
截尾	16	13.01	0.642	4.7	<0.001
汉语					
AW/VW	9	11.42	0.690		
截尾	16	12.86	0.662	4.7	<0.001
VW/AW	13	11.42	0.690	2.3	≈0.05
截尾	17	12.86	0.662	6.3	≈0.000 1

在汉语的实验中,最好的结果是比较八个词的听觉呈现与视觉呈现,把视觉汉字(VW)作为原型,听觉汉字(AW)作为测试样本,在截尾的情况下得到了从二十四个试样中正确地分类出十七个的最好结果,p 约等于 0.000 1,并且像前面一样,重复次数 r 在零假设条件下大于一百万。因此,我们以强化了贝克莱和休谟实验结果而告终,没有断言,这个证据迄今为止是完全

决定性的。

适当地停止这种分析之处是强调,截尾不能确保改进分类。465
表6中所示的许多其他结果在进行数据截尾后并没有大的改进,但另一方面,对于那里报告的实验来说,除了视觉图像实验之外,这些结果在数据截尾前是很显著的。

EEG记录,像在许多科学领域中一样,用于排除数据中的伪差和其他反常的统计方法和实验方法,构成了具有复杂文献的一个大主题。我这里只报告普通的统计方法,但是,我很高兴以"清理"数据的典型方法的这个例子来结束。正是应该用这样的截尾数据形成一种表征,这种表征适用于检验某个理论,或通常情况下,是检验许多理论思想。[1]

8.7 结语:科学中的表征与还原[2]

还原的论题与哲学一样古老。早在公元前5世纪,留基伯(Leucippus)特别是他的学生德谟克利特,就曾用明确的一般术语概括陈述了著名的原子论思想。物质只不过是在空的空间中运动的坚硬的原子。一个物体的重量、密度和硬度只依赖于它的原子在像希腊人所描述的空的空间那样的真空中的排列。而且,按照德谟克利特的观点,除了遵照定律发生和必然地发生的事件之外,没有什么是随机发生的。尽管几个世纪以来,这种还原论的机械自然观受到了亚里士多德及其追随者的许多批评,

---

① 原文"congerie"有误,应该为"congeries"。——译者

② 本节的这些评论是对苏佩斯著作中(1984, pp. 120 - 125)的那些关于还原论的评论的补充。

但是,在 17 和 18 世纪,这种观点再一次对物理学或当时著名的自然哲学的发展起到了积极作用。

不过,德谟克利特之后出现的两位最著名的古代原子论者为这种机械论的图像增添了偶然的效果。下面是伊壁鸠鲁的一段著名引文,节选自他给希罗多德的信,此信大概写于公元前 4 世纪(Oates 1940, p. 5):

> 原子总是不断地运动的,有些直接落下,有些转弯,有些经过碰撞后弹回。在后一种情况下,有些彼此离得很远,而有些每当它们碰巧被交错的其他原子阻止时,就再次弹来弹去,否则,就被它们周围交错的原子夹住。因为一方面,把每个原子分离开来的空间本性导致了这一点,就像不能经得起阻力一样;另一方面,这些原子的硬度使它们在碰撞后弹回来,交织后弹回的距离与碰撞后允许分开的距离一样。而且,这些运动没有开始,因为这些原子和空间就是原因。

466

下面是卢克莱修在写于公元前 1 世纪的《物性论》(*De Rerum Natura*)中不得不指出的(Oates 1940, p. 95):

> 这里的这种观点是我们希望你理解的:当物体由于它们自身的重量完全穿越真空落下时,它们在不确定的时间和地点会有点偏离它们的运动轨道:你只能称之为改变倾向。如果它们不转弯,那么它们全部像雨点一样落入很深的真空,而且,在一开始,既不会产生对撞,也不会产生猛击:因此,自然界从来没有产生出什么。

这样,伊壁鸠鲁和卢克莱修提出把自然界还原为概率力学,一门直到 20 世纪才在科学上得到真正发展的学科。当然,除了一些著名的例外,从亚里士多德到牛顿,这些原子论者并没有受到好

评。在牛顿于 1692/3 年 2 月 25 日写给理查德·本特利(Richard Bentley)的一封著名的信中,刚好找到了下列一段引文。

> 如果在伊壁鸠鲁的意义上,万有引力在物质中是基本的和固有的,那么,无生命的物质,在没有其他非物质的东西为媒介的条件下,像它一定所是的那样,没有相互接触,就应该影响和作用于别的物质,这是不可思议的。这就是为什么我希望你不要把天赋的重力归功于我的一个理由所在。对于物质来说,重力应该是天赋的、固有的和基本的,以便一个物体可以通过真空超距地作用于另一个物体,在没有任何其他东西为媒介的前提下,据此,物体的作用和力可以从一个物体传递到另一个物体,这一点对我来说是很谬论的,以至于我相信,有足够哲学思维能力的人,都不会沉溺于此。按照某些定律,重力一定是由一个不变的媒介作用引起的;但不管这种媒介是物质的,还是非物质的,我都留给我的读者去考虑。

自牛顿以来,对物理学的发展做出伟大贡献的人几乎毫无例外地对在科学上(与关于还原的推测思想相反)确立的还原论题之间的差别相当清楚。牛顿之后大约两百年,麦克斯韦在他的《物质与运动》一本小册子中,非常明确地陈述了这种观点。

> 当然,如果我们通过假设,真正的物体是由物质构成的系统(这与我们已经给出的定义在所有方面都一致)开始,那么我们可以继续断言,所有的现象都是位型和运动的变化,尽管我们不准备定义用来说明特殊现象的这种位型和运动。但在精密科学中,对这样断言的说明必须根据它们的性能,而不是根据它们的诺言作出评价。一个系统的位

型和运动是能以一种精确方式加以描述的事实，因此，为了可以把通过一个物体系统的位型和运动对一种现象的说明看成是增加了我们的科学知识，位型、运动和力必须得到明确的阐述，并表明与已知的事实以及说明现象的能力相一致。

467

(Maxwell, *Matter and Motion*, 1877/1920: pp. 70 - 71)

麦克斯韦很清楚地解释了把坚定地确立的还原和不是这样确立的还原区分开来的必要。这并不意味着，所有的推测都是一件坏事。我们甚至坚持认为，在科学的大多数情况下，在“下一个大的理论”提出之前，先是通过推测确定，这个大的理论是有待发现的或发明的。就近代的本质而言，重要的是，科学推测被看成是直接的猜想，不能混同于也不能当作是含有形而上学色彩的可疑成分。本书严谨的论题之一是，科学还原是一件难事。用详细的和明确的表征定理来完成科学还原更加困难。尽管本书的许多页一直关注表征定理，但是，我一直没有证明乃至阐明实际上有意义的这样一个表征定理：它能为把科学的一部分还原为另一部分的某种重要主张提供坚实的支持。错过我前面对这个主题的评论的读者可能会问，为什么没有呢？

答案是显而易见的。这个问题对我来说太难。也许，要考虑的最著名的案例是把热力学还原为统计力学。在一般的概念层次上，这是过去常常被称为机械论哲学的伟大胜利之一。并且在一般分析层次上，恩斯特·内格尔于 1949 年发表了关于这个主题的一篇著名文章。但内格尔完全意识到，他不能为一个适当的表征定理给出完全详尽的或开始提供严格的数学证明。吕埃勒(1978)给出的一条严格的数学进路是最好的事例之一，但是，为了显示明显的相变，他的理想化要求直接应用的物理系

统，实际上，在所有方向上，即在充满的整个空间里，是极大的。在宏观热力学系统中对相变的详细说明是经典平衡统计力学的主要目标之一。我也想强调，在我们自然希望说明的自然现象中，相变在最明确的系统变化之中。例如，我们喜欢从基本的物理学原理推出这种结果：在常压下，液态水在大约摄氏一百度时转变为水蒸气。从厨房里的茶壶的基本原理推出这个结果仍然是不可能的。其他戏剧性的相变也一样，例如，一杯水的冰冻。

这些引人注目的日常现象缺乏严格的表征定理，并不意味着物理学家是在进行坏的分析工作。这恰好是我们开始以明确的方式意识到，为所有类型的日常自然现象提供严格的数学结果是多么困难。可是物理学家，即解决问题的聪明人，已经率先了解到许多这样的现象，包括在没有进行严格辩护的条件下如何计算大量的定量结果在内。

根据从天气预报到股票市场范围内的例子，已知有多少这种情况推广到其他的重要现象，像在其他问题上一样，哲学上重要的是，对所能实现的东西采取多元论的进路。现在的天气预报比四十年前好了许多，即使我们现在意识到，提前一个月作出详细的天气预报，在我们当前的理论和分析天气数据的计算体系中是没有希望的。物理学家善于推导出有用的和高度精确的方程来说明从一般数学的观点看无法严格分析的许多现象。我们从这样的工作获得了关于什么样的表征和还原可能是近似正确的见识与直觉，即使在至今未知的小路上，也可能潜藏着新的意外。不过，对于科学哲学家来说，重要的是对物理学家当前认为是真的许多物理现象，或诸如市场行为之类的其他现象，或别的科学家所用的语言，作出理解和概述。事实上，在许多科学领域内，关于表征的必然的主要的经验结果可能是当前所能获得 468

的惟一结果。第 6 节报告的语言的脑电图表征就是一个例子。

但是,为关于表征和还原的严格结果留有余地,也是这样一种多元论观点的应有之意。在后面的四章中,我已经试图在不同的科学领域内举了一些足够简单的易操作的事例。我在这些事例的列表中包括了第 5 章对概率的表征或解释的扩展分析。同样明显的是,并非所有的表征定理都直接地推崇还原。例如,在测量理论中可证明的许多表征定理(我在第 3 章给出的基本例子)没有在概念上建议,我们实质上正在把物理特性或心理特性还原为数,而是说用数来提炼我们对这些特性的分析。换言之,这种表征不同于在物理学上把温度概念还原为粒子的平均运动。

另一方面,像在第 3 章中陈述的那样,通过寄存器机或图灵机对可计算函数的表征,或者,在第 8 章中通过刺激-反应模型对有穷自动机或寄存器机的表征,确实建议对可计算过程的相当直接的物理实现,而且,这些实现实际上比早在一百年前想到的机制更简单得多。对于许多复杂现象——因此而在一种意义上对他们进行的还原——来说,这些简单机制的存在性是在最近很厚的一本巨著《一种新的科学》(2002)中多次强调与举例说明的一种观点,这本著作的作者是理论物理学家斯蒂芬·沃尔弗拉姆(Stephen Wolfram)。但沃尔弗拉姆像他在早期作品中那样,正确地强调了许多这样的过程在可计算的意义上是不可还原的,意指根本没有比观察这些过程本身更简单的方式来研究它们的行为。

如果不讨论新技术的复杂性,显然,像计算的不可还原性那样的概念,在有助于说明在我们描述的日常现象中为什么公认的还原案例没有导致概念的剧变时,是有用的。当面对像是一场概念革命时,为什么只有很小的变化,当然有其他理由。我相

信，我们在未来一百年内还要像现在一样谈论桌椅。而对于不计其数的其他熟悉的对象和过程来说，也是如此。

另一方面，这样的还原并不影响日常思想和讨论，持有这种论点也是错误的。现在，老百姓以各种方式使用对时间和温度的数值表征，在不太遥远的过去，这似乎是奇怪的和不恰当的。469 如同我们在更多的物理细节上开始了解我们的身心是如何工作的一样，有某些类似的变化似乎是可能的，但是，未来的表征及其不变性将成为新的自然的习惯用语的基础，这种预测，即使在可计算的意义上是可还原的，也难以尝试。

附录

471 各章表征和不变性定理汇总表[①]

3 同构表征理论

表 征 定 理

3.1 有穷弱排序

3.2 无穷弱排序

4.1 外延测量

4.2 差测量

4.3 对分测量

4.4 联合测量

5.1 通过无限寄存器机表征的部分递归函数

不 变 性 定 理

无

4 不变性

表 征 定 理

无

不 变 性 定 理

3.1 定序测量

3.2　外延测量

3.3　对分测量

3.4　联合测量

5.4　各态历经的马尔可夫过程的熵

5　概率的表征　472

表 征 定 理

2.1　有穷拉普拉斯概率空间

3.1　结果的有穷序列

3.3　结果的无穷序列

3.6　丘奇意义上的随机无穷序列

4.4　柯尔莫哥洛夫意义上的复杂的有穷序列

5.1　卡尔纳普的正则测度

5.3　丘瓦基的随机过程的拉普拉斯概率空间

6.1　标准序列

6.2　衰变倾向

6.3　反应强度倾向

6.4　掷硬币时出现正面的倾向

6.5　三体问题中的随机倾向

7.3　斯科特定理

7.4　定性的期望值

7.4　定性的概率

7.5　定性的条件期望值

7.5　定性的条件概率

---

①　定理按节来编号，因此，在一节中是连续的。这样，3.1 是第 3 节的定理 1。前面两章中没有定理。

7.6　可交换的序列(德·菲内蒂)

7.7　定性的近似相信的测量

不变性定理

7.4　定性的期望

7.4　定性的概率

7.5　定性的条件期望

7.5　定性的条件概率

7.7　定性的接近相信的测量

6　空间与时间的表征

表征定理

1.1　仿射空间

1.3　欧几里得空间

6.1　限于全等的视觉空间

8.5　有穷仿射结构

不变性定理

1.2　仿射空间

1.4　欧几里得空间

2.1　经典时空

3.1　相对论性的时空

6.1　限于全等的视觉空间

6.6　有穷仿射结构

473

7　力学中的表征

表征定理

1.4　系统的质心

1.5　系统的角动量的变化率

1.7 嵌入粒子系统

1.9 势能函数

2.1 定域隐变量

2.2 对称的定域隐变量

2.3 满足二阶因式分解的定域隐变量

2.4 满足定域条件Ⅱ的隐变量

2.5 三个随机变量的不等式

2.6 贝尔不等式

2.7 GHZ的理想测量

2.8 GHZ实验的不等式

2.9 联合高斯分布

2.10 缺少协方差的联合高斯分布

2.11 三个高斯随机变量的不等式

不变性定理

1.6 动量守恒与角动量守恒

1.8 能量守恒

3.1 马尔可夫特性的时间不变性

3.2 强可逆的马尔可夫链的时间不变性

3.3 强可逆的连续时间马尔可夫链的时间不变性

3.4 二阶马尔可夫链的时间不变性

8 语言的表征

表征定理

1.2 语境无关文法的范式(乔姆斯基)

1.3 语境无关文法的范式(格雷巴赫)

2.6 正则文法

2.7 有穷自动机

2.8 在一定特性条件下(克伦)闭合的正则语言

2.10 语境无关文法

2.11 下推自动机

2.12 语境敏感文法

2.12 线性有界自动机

2.13 短语结构语法

3.1 关于有穷自动机的刺激-反应模型定理

3.2 关于正则语言的刺激-反应模型定理

3.3 米勒和乔姆斯基意义上的 Tote 层次结构的刺激-反应模型定理

3.4 关于感知呈现的有穷自动机定理和刺激-反应模型定理

3.5 关于部分递归函数的寄存器学习模型定理

4.13 关于线性学习模型的刺激-反应模型定理

不变性定理

无

参考文献

Abelson, R. M. and R. A. Bradley (1954). A 2x2 factorial with paired comparisons. *Biometrics* 10(4):487–502.

Aczél, J. (1961). Über die Begründung der Additions- und Multiplikationsformeln von bedingten Wahrscheinlichkeiten. *Magyar Tudományos Akadémia Matematikai Kutató Int. Közleményei* 6:110–122.

Aczél, J. (1966). *Lectures on functional equations and their applications*. New York: Academic Press.

Aho, A. V. and J. D. Ullman (1972). *The theory of parsing, translation, and compiling*, vol. 1. Englewood Cliffs, NJ: Prentice Hall.

Alekseev, V. M. (1968–1969*a*). Quasirandom dynamical systems I, II, III. *Math. USSR Sbornik* 5,6,7:505–560 and 1–43.

Alekseev, V. M. (1969*b*). Quasirandom dynamical systems. (Russian). *Mat. Zametki* 6(4):489–498. English translation in Math. Notes 6 (1969), 749–753.

Alexandrov, A. D. (1950). On Lorentz transformations. (Russian). *Uspekhi Matematichekikh Nauk* 5, 3(37):187.

Alexandrov, A. D. (1975). Mapping of spaces with families of cones and space-time transformations. *Annali di Matematica Pura ed Applicata* 103:229–257.

Ames Jr., A. (1946). Binocular vision as affected by relations between uniocular stimulus patterns in common place environments. *American Journal of Psychology* 59:333.

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Lawrence Erlbaum.

Angell, R. B. (1974). The geometry of visibles. *Noûs* 8:87–117.

Appell, P. (1941). *Traité de mécanique rationelle. Tome I: Statique, dynamique du point*. Paris: Gauthier-Villars, sixth edn.

Aquinas, T. (1944). Summa theologica. In A. C. Regis, ed., *Basic writings of Saint Thomas Aquinas*, vol. 1. New York: Random House.

Arbib, M. A. (1969). Memory limitations of stimulus-response models. *Psychological Review* 76:507–510.

Archimedes (1897). On the equilibrium of planes. In T. L. Heath, ed., *The works of Archimedes, edited in modern notation with introductory chapters*, pages 189–220. Cambridge, England: Cambridge University Press. Translation by T. L. Heath.

Aristotle (1960). *De Caelo (On the heavens)*. Cambridge, MA: Harvard University Press. English translation by W. K. C. Guthrie. First published in 1939.

Aristotle (1975). *De Anima (On the soul)*. Cambridge, MA: Harvard University Press, 4th edn. English translation by W. S. Hett. First published in 1936.

Arnold, V. I. (1978). *Mathematical methods of classical mechanics*. New York: Springer.

Arrow, K. J. (1951). *Social choice and individual values*. New York: Wiley.

Attias, H. (1999). Independent factor analysis. *Neural Computation* 11:803–852.

Balzer, W. (1978). *Empirische Geometrie und Raum-Zeit-Theorie in Mengentheoretischer Darstellung*. Königstein/Ts.: Scriptor Verlag.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society* 53:370–418. Reprinted in *Facsimiles of two papers by Bayes* with a commentary by W. Edwards Deming, New York, Hafner Pub. Co., 1940.

Beck, A. (1890). Die Ströme der Nervencentren. *Centralblatt für Physiologie* 4:572–573.

Bell, J. S. (1964). On the Einstein-Podolsky-Rosen paradox. *Physics* 1:195–200.

Bentham, J. (1789/1907). *The principles of morals and legislation*. Oxford: Clarendon Press. First published in 1789.

Berger, H. (1929). Uber das Elektroenkephalogram des Menschen. *Archiv für Psychiatrie Nervenkrankheiten* 87(4):527–570. Translated by P. Gloor in *Electroencephalography Clinic Neurophysiology*, Supplement **28**, 37–73, 1969.

Berkeley, G. (1709/1901). An essay towards a new theory of vision. In A. C. Fraser, ed., *Berkeley's complete works*, vol. 1, pages 93–210. New York: Oxford University Press. First published in 1709.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Comentarii academiae scientiarum imperiales petropolitanae* 5:175–192. Translation by L. Sommer in *Econometrica* 1954, **22**, 23–26.

Bernoulli, J. (1713). *Ars conjectandi*. Basileae: Impensis Thurnisiorum fratrum. Reprinted on pp. 106-286 in Vol. 3 of *Die Werke von Jakob Bernoulli*, Birkhäuser Verlag, Basel, 1975.

Bertrand, J. L. (1888). *Calcul des probabilitiés*. Paris: Gauthier-Villars.

Bever, T. G., J. Fodor, and M. Garrett (1968). A formal limitation of associationism. In T. R. Dixon and D. L. Horton, eds., *Verbal behavior and general behavior theory*. Englewood Cliffs: Prentice-Hall.

Blank, A. A. (1953). The Luneburg theory of binocular visual space. *Journal of the Optical Society of America* 43:717–727.

Blank, A. A. (1957). The geometry of vision. *British Journal of Physiological Optics* 14:154–169.

Blank, A. A. (1958*a*). Analysis of experiments in binocular space perception. *Journal of the Optical Society of America* 48:911–925.

Blank, A. A. (1958*b*). Axiomatics of binocular vision: The foundations of metric geometry in relation to space perception. *Journal of the Optical Society of America* 48:328–334.

Blank, A. A. (1961). Curvature of binocular visual space: An experiment. *Journal of the Optical Society of America* 51:335–339.

Blumenfeld, W. (1913). Untersuchungen über die scheinbare Grösse in Sehraume. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* 65:241–404.

Bohm, D. (1965). *The special theory of relativity*. New York: W. A. Benjamin.

Borsuk, K. and W. Szmielew (1960). *Foundations of geometry*. Amsterdam: North Holland.

Bourbaki, N. (1950). The architecture of mathematics. *American Mathematical Monthly* 57:231–232.

Bouwmeester, D., J. W. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger (1999). Observation of three-photon Greenberger-Horne-Zeilinger entanglement. *Physical Review Letters* 82(7):1345–1349.

Bowerman, M. (1989). Learning a semantic system. What role do cognitive predispositions play? In M. L. Rice and R. L. Schiefelbusch, eds., *The teachability of language*, pages 133–169. Baltimore: Paul H. Brookes.

Bowerman, M. (1996). The origin of children's spatial semantic categories: Cognitive vs. linguistic determinants. In J. J. Gumperz and S. C. Levinson, eds., *Rethinking linguistic relativity*, pages 145–176. Cambridge, England: Cambridge University Press.

Bradley, R. A. (1954*a*). Rank analysis of incomplete block designs, II. Additional tables for the method of paired comparisons. *Biometrika* 41:502–537.

Bradley, R. A. (1954*b*). Incomplete block rank analysis: On the appropriateness of the model for a method of paired comparisons. *Biometrics* 10:375–390.

Bradley, R. A. (1955). Rank analysis of incomplete block designs, III. Some large-sample results on estimation and power for a method of paired comparisons. *Biometrika* 42:450–470.

Bradley, R. A. and M. E. Terry (1952). Rank analysis of incomplete block designs, I. The method of paired comparisons. *Biometrika* 39:324–345.

Breiman, L. (1999). Prediction games and arcing algorithms. *Neural Computation* 11:1493–1518.

Bresnan, J. (1982). The passive in lexical theory. In J. Bresnan, ed., *The mental representation of grammatical relations*, pages 3–86. Cambridge, MA: MIT Press.

Brontë, C. (1847/1971). *Jane Eyre, an authoritative text, backgrounds, criticism, edited by Richard J. Dunn*. New York: W. W. Norton & Company. First published in 1847.

Bruns, H. (1887). Über die Integrale des Vielkörper-Problems. *Acta Mathematica* 11:25–96.

Brush, S. G. (1976). *The kind of motion we call heat: A history of the kinetic theory of gases in the 19th century*. Amsterdam: North-Holland Publishing Co.

Buffon, G. L. (1733). Communication. *Actes de l'académie des sciences - Paris* pages 43–45.

Buffon, G. L. (1777). *Essai d'arithmétique morale*.

Busemann, H. (1955). *The geometry of geodesics*. New York: Academic Press.

Bush, R. R. and W. K. Estes, eds. (1959). *Studies in mathematical learning theory*. Stanford, CA: Stanford University Press.

Bush, R. R. and F. Mosteller (1955). *Stochastic models for learning*. New York: Wiley.

Campbell, N. R. (1920). *Physics, the elements*. Cambridge University Press.

Campbell, N. R. (1928). *An account of the principles of measurement and calculation*. New York: Longmans, Green and Co.

Cantor, G. (1895). Beiträge zur Begründung der transfiniten Mengenlehre. *Mathematische Annalen* 46:481–512. Translation by P. E. B. Jourdain in *Contributions to the founding of the theory of transfinite numbers*, Dover, New York 1952, pp. 85–136 and pp. 137–201, with original pagination indicated.

Carnap, R. (1945). On inductive logic. *Philos. Sci.* 12:72–97.

Carnap, R. (1950). *The logical foundations of probability*. Chicago, IL: University of Chicago Press.

Carnap, R. (1952). *The continuum of inductive methods*. Chicago, IL: The University of Chicago Press.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of Science*. Cambridge, England: Cambridge University Press.

Caton, R. (1875). The electric currents of the brain. *British Medical Journal* 2:278. Abstract only.

Chaitin, G. J. (1966). On the length of programs for computing finite binary sequences. *Journal of the ACM* 13(4):547–569.

Chaitin, G. J. (1969). On the length of programs for computing finite binary sequences. *Journal of the ACM* 16(1):145–159.

Choi, S. and M. Bowerman (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41(1):83–121.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control* 2:137–167.

Chuaqui, R. (1977). A semantical definition of probability. In A. I. Arruda and N. C. A. da Costa, eds., *Non classical logic, model theory and computability*, pages 135–168. Amsterdam: North Holland.

Chuaqui, R. (1991). *Truth, possibility, and probability*. Amsterdam: North Holland.

Chuaqui, R. and P. Suppes (1990). An equational deductive system for the differential and integral calculus. In P. Martin-Löf and G. Mints, eds., *Lecture notes in computer science, Proceedings of COLOG-88 International Conference on Computer Logic*, pages 25–49. Berlin: Springer.

Chuaqui, R. and P. Suppes (1995). Free-variable axiomatic foundations of infinitesimal analysis: A fragment with finitary consistency proof. *The Journal of Symbolic Logic* 60:122–159.

Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics* 58:345–363.

Church, A. (1940). On the concept of a random sequence. *Bulletin of American Mathematical Society* 46(2):130–135.

Clagett, M. (1959). *The science of mechanics in the middle ages*. Madison, WI: University of Wisconsin Press.

Compton, A. H. (1923). A quantum theory of the scattering of X-rays by light elements. *Physical Review* 21:483–502.

Cooley, J. W. and J. W. Tukey (1965). An algorithm for the machine computation of complex Fourier series. *Math. Computation* 19:297–301.

Copeland, A. H. (1928). Admissible numbers in the theory of probability. *Amer. Jour. Math.* 50(4):535–552.

Coren, S. and J. S. Girgus (1978). *Seeing is deceiving : the psychology of visual illusions*. Hillsdale, NJ: Lawrence Erlbaum.

Cover, T. M., P. Gacs, and R. M. Gray (1989). Kolmogorov's contribution to information theory and algorithmic complexity. *Annals of Probability* 17:840–863.

Cover, T. M. and J. A. Thomas (1991). *Elements of information theory*. New York: Wiley.

Crangle, C. and P. Suppes (1989). Geometrical semantics for spatial prepositions. In P. A. French, T. E. Uehling, Jr., and H. K. Wettstein, eds., *Midwest studies in philosophy*, vol. 14, pages 399–422. Notre Dame, IN: University of Notre Dame Press.

da Costa, N. C. A. and R. Chuaqui (1988). On Suppes' set theoretical predicates. *Erkenntnis* 29:529–589.

Daniels, N. (1972). Thomas Reid's discovery of a non-Euclidean geometry. *Philosophy of Science* 39:219–234.

David, F. N. (1962). *Games, gods and gambling: The origins and history of probability and statistical ideas from the earliest times to the Newtonian era*. London: Charles Griffin.

Davidson, D. and P. Suppes (1956). A finitistic axiomatization of subjective probability and utility. *Econometrica* 24:264–275.

Davidson, D., P. Suppes, and S. Siegel (1957). *Decision making: An experimental approach*. Stanford, CA: Stanford University Press. Midwahy Reprint, Chicago, IL: University of Chicago, Press, 1977.

Davis, M. (1958). *Computability and unsolvability*. New York: McGraw–Hill.

Davis, R. L. (1953). The number of structures of finite relations. *Proceedings of the American Mathematical Society* 4:486–495.

Dawid, A. P. (1986). Probability forecasting. In S. Kotz, N. L. Johnson, and C. B. Read, eds., *Encyclopedia of statistical sciences*, vol. 7, pages 210–218. New York: Wiley.

de Barros, A. J. and P. Suppes (2000). Inequalities for dealing with detector inefficiencies in Greenberger-Horne-Zeilinger-type experiments. *Physical Review Letters* 84:793–797.

de Finetti, B. (1931). Sul significato della probabilità. *Fundamenta Mathematicae* 17:298–329.

de Finetti, B. (1937/1964). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré* 7:1–68. Translation in Kyburg and Smokler (1964), 93–158.

de Finetti, B. (1951). Recent suggestions for the reconciliation of theories of probability. In J. Neyman, ed., *Proceedings of the second Berkeley symposium on mathematical statistics and probability*. Berkeley: University of California Press.

de Finetti, B. (1974). *Theory of probability*, vol. 1. New York: Wiley. Translated by A. Machi and A. Smith.

de Finetti, B. (1975). *Theory of probability*, vol. 2. New York: Wiley. Translated by A. Machi and A. Smith.

de Moivre, A. (1718/1756). *The doctrine of chances*. London: Chelsea, 3rd edn. First published in 1718.

Dedekind, R. (1872). *Stetigkeit und Irrationale Zahlen*. Vieweg: Braunschweig. Translated by W. W. Beman, Continuity and irrational numbers, in 'Essays on the theory of numbers', 1-27, New York: Dover, 1963.

Dembowski, P. (1968). *Finite geometries*. New York: Springer.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statist.* 38:325–340.

Descartes, R. (1637/1954). *La Géométrie*. New York: Dover. Translated from the French and Latin by D. E. Smith and M. L. Latham. First published in 1637.

Descartes, R. (1644/1983). *Principia philosophiae*. Amsterdam: Ludovicum Elzevirium. Translation by V. R. Miller and R. P. Miller, *Principles of philosophy*, Reidel, Dordrecht, 1983. First published 1644.

Descartes, R. (1649/1927). Passions of the soul. In R. M. Eaton, ed., *Descartes selections*, pages 361–403. New York: Charles Scribner's sons. First published in 1649.

Diaconis, P. (1977). Finite forms of de Finetti's theorem on exchangeability. *Synthese* 36:271–281.

Diaconis, P. and D. Freedman (1980). Finite exchangeable sequences. *The Annals of Probability* 8(4):745–764.

Dijksterhius, E. J. (1957). *Archimedes*. New York: Humanities Press.

Dirac, P. A. M. (1947). *The principles of quantum mechanics*. Oxford: Clarendon Press, 3rd edn. First published in 1930.

Domotor, Z. (1969). Probabilistic relational structures and their applications. Tech. Rep. 144, Stanford University, Institute for Mathematical Studies in the Social Sciences, Stanford, CA.

Domotor, Z. (1972). Causal models and space-time geometries. *Synthese* 24:5–57.

Doob, J. L. (1941). Probability as measure. *Ann. Math. Stat.* 12:206–214, 216–217.

Doob, J. L. (1960). Some problems concerning the consistency of mathematical models. In R. E. Machol, ed., *Information and Decision Processes*, pages 27–33. New York: McGraw–Hill.

Drösler, J. (1966). Das beidaeugige Raumsehen. In W. Metzger, ed., *Handbuch der Psychologie*, pages 590–615. Goettingen: Hogrefe.

Drösler, J. (1979*a*). Foundations of multi-dimensional metric scaling in Cayley-Klein geometries. *British Journal of Mathematical and Statistical Psychology* 32:185–211.

Drösler, J. (1979*b*). Grundlegung einer mehrdimensionalen metrischen Skalierung. *Zeitschrift fuer experimentelle und angewandte Psychologie* 26:1–36.

Drösler, J. (1979*c*). Relativistic effects in visual perception of real and apparent motion. *Arch. Psychol.* 131:249–266.

Drösler, J. (1987). Experimentelle untersuchung der Geometrie des monokularen Sehraumes. *Zeitschrift fuer experimentelle und angewandte Psychologie* 3:351–369.

Drösler, J. (1988). The psychophysical function of binocular space perception. *Journal of Mathematical Psychology* 32:285–297.

Drösler, J. (1992). Eine untersuchung des perspektivischen Sehens. *Zeitschrift fuer experimentelle und angewandte Psychologie* 34:515–532.

Du Bois-Reymond, E. (1848). *Untersuchungen über thierische Elektricität*. Berlin: Verlang von G. Reiner. Passage quoted translated by Hebbel E. Hoff, Galvani and the pre-Galvanic electrophysiologists, *Annals of Science*, 1 (1936), 157–172.

Einstein, A. (1905/1923). Zur electrodynamik bewegter Körper. *Annalen der Physik* 17. Translation in Lorentz et al. (1923).

Einstein, A. (1934). On the method of theoretical physics. *Philosophy of Science* 1:163–169.

Einstein, A. (1949). Remarks to the essays appearing in this collective volume. In P. A. Schilpp, ed., *Albert Einstein: Philosopher-scientist*, Library of Living Philosophers, Vol. 7, pages 663–688. Chicago: Open Court Publishing.

Epelboim, J. and P. Suppes (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research* 41:1561–1574.

Epicurus (1940). Epicurus to Herodotus. In Oates (1940), pages 3–15.

Estes, W. K. (1959*a*). Component and pattern models with Markovian interpretations. In Bush and Estes (1959).

Estes, W. K. (1959*b*). The statistical approach to learning theory. In S. Koch, ed., *Psychology: A Study of a Science*. New York: McGraw-Hill.

Estes, W. K. and P. Suppes (1959*a*). Foundations of linear models. In Bush and Estes (1959), pages 137–179.

Estes, W. K. and P. Suppes (1959*b*). Foundations of statistical learning theory, II. The stimulus sampling model. Tech. Rep. 26, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Estes, W. K. and P. Suppes (1974). Foundations of stimulus sampling theory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes, eds., *Contemporary developments in mathematical psychology, Vol. 1: Learning, memory and thinking*, pages 163–183. San Francisco, CA: Freeman.

Euclid (1926). *Elements. The thirteen books of Euclid's Elements*. New York: Dover, 2nd edn. Translated from the text of Heiberg, with introduction and commentary, by Sir Thomas L. Heath.

Euclid (1945). Optics. *Journal of the Optical Society of America* 35. Translated by H. E. Burton.

Euclid (1996). *Euclid's Phaenomena: A translation and study of a Hellenistic treatise in spherical astronomy*. New York: Garland Publishing. Translated by J. L. Berggren and R. S. D. Thomas.

Euler, L. (1842). *Lettres de Euler à une Princesse d'Allemagne, sur divers sujets de physique et de philosophie*. Paris: L. Hachette.

Fagin, R. (1976). Probabilities on finite models. *Journal of Symbolic Logic* 41:50–58.

Feldman, J. A., G. Lakoff, D. Bailey, S. Narayanan, T. Regier, and A. Stolcke (1996). L0: The first five years. *Artificial Intelligence Review* 10:103–129.

Feldman, J. A., G. Lakoff, A. Stolcke, and S. Hollbach Weber (1990). Miniature language acquisition: A touchstone for cognitive science. In *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, pages 686–693. Cambridge, MA: MIT Press.

Feller, W. (1950). *An introduction to probability theory and its applications*, vol. 1. New York: Wiley.

Feller, W. (1966). *An introduction to probability theory and its applications*, vol. 2. New York: Wiley.

Fine, A. (1982). Hidden variables, joint probability, and the Bell inequalities. *Physical Review Letters* 48:291–295.

Fine, T. L. (1973). *Theories of probability*. New York: Academic Press.

Finke, R. A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT Press.

Fishburn, P. C. (1970). *Utility theory for decision making*. New York: Wiley.

Fitelson, B. (2001). *Studies in Bayesian Confirmation Theory*. Ph.D. thesis, University of Wisconsin-Madison, Ann Arbor, MI.

FitzGerald, G. F. (1889). The ether and the Earth's atmosphere. *Science* 13(328):390.

Fock, V. A. (1931/1978). *Fundamentals of quantum mechanics*. Moscow: Mir Publishers. Translated from the Russian by E. Yankovsky.

Fodor, J. and Z. Pylyshyn (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition* 28:3–71.

Foley, J. M. (1964). Desarguesian property in visual space. *Journal of the Optical Society of America* 54:5:684–692.

Foley, J. M. (1965). Visual space: A scale of perceived relative direction. *Proceedings of the 73rd Annual Covention of the American Psychological Association* 1:49–50.

Foley, J. M. (1966). Locus of perceived equidistance as a function of viewing distance. *Journal of the Optical Society of America* 56:822–827.

Foley, J. M. (1969). Distance in stereoscopic vision: The three-point problem. *Vision Research* 9:1505–1521.

Foley, J. M. (1972). The size-distance relation and intrinsic geometry of visual space: Implications for processing. *Vision Research* 12:323–332.

Foley, J. M. (1978). Distance perception. In R. Held, H. Leibowitz, and H. L. Teuber, eds., *Handbook of sensory physiology, Vol. 8: Perception*, pages 181–213. New York: Springer.

Ford, L. R., Jr. (1957). Solution of a ranking problem from binary comparisons. *American Mathematical Monthly* pages 28–33.

Frank, P. (1909). Die Stellung des Relativitätsprinzips in System der Mechanik und der Elektrodynamik. *Sitzungsberichte der Akademie der Wissenschaften in Wien. Mathemat. Naturw. Klasse* 118(2a):373–446.

Frege, G. (1899). Letter to Hilbert. In Frege (1980), pages 35–36.

Frege, G. (1900). Letter to Hilbert. In Frege (1980), page 43.

Frege, G. (1903*a*). Über die Grundlagen der Geometrie. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 12:319–324. Translation in Frege (1971).

Frege, G. (1903*b*). Über die Grundlagen der Geometrie II. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 12:368–375. Translation in Frege (1971).

Frege, G. (1906*a*). Über die Grundlagen der Geometrie I. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15:293–309. Translation in Frege (1971).

Frege, G. (1906*b*). Über die Grundlagen der Geometrie (Fortsetzung) II. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15:377–403. Translation in Frege (1971).

Frege, G. (1906*c*). Über die Grundlagen der Geometrie (Schluss) III. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15:423–430. Translation in Frege (1971).

Frege, G. (1971). *On the foundations of geometry and formal theories of arithmetic*. London and New Haven: Yale University Press. Translation by E. H. Kluge.

Frege, G. (1980). *Philosophical and mathematical correspondence*. Chicago, IL: University of Chicago Press. Edited by G. Gabriel.

Freudenthal, H. (1965). Lie groups in the foundations of geometry. *Advances in Mathematics* 1:145–190.

Friedman, J., T. Hastie, and R. Tibshirani (2000). Additive logistic regression: A statistical view of boosting. *Annals of Statistics* 28:337–407.

Fu, K. S. (1974). *Syntactic methods in pattern recognition*. New York: Academic Press.

Gaifman, H. (1964). Concerning measures in first order calculi. *Israel J. Math.* 2:1–18.

Gaifman, H. and M. Snir (1982). Probabilities of rich languages, testing and randomness. *Journal of Symbolic Logic* 47:495–548.

Galavotti, M. C. (1987). Comments on Patrick Suppes "Propensity interpretations of probability". *Erkenntnis* 26:359–368.

Galavotti, M. C. (In press). Between logicism and subjectivism: Harold Jeffreys's probabilistic epistemology. *British Journal for the Philosophy of Science.*

Galison, P. (1987). *How experiments end*. Chicago: University of Chicago Press.

Galvani, L. (1791/1953). De viribus electricitatis in motu musculari. *De Bononiensi Scientiarum et Artium Instituto atque Academia, Comm.* 7:363–418. Translated by Margaret Glover Foley, in *Luigi Galvani, Commentary on the Effects of Electricity on Muscular Motion.* Notes and introduction by I. Bernard Cohen. Norwalk, CT: Burndy Library, 1953.

Galvani, L. (1794). *Dell'uso, e dell'attivitá, dell'arco conduttore nelle contrazioni dei muscoli.* Bologna: A. S. Tommaso d'Aquino. Published anonymously. Portion translated in: B. Dibner, *Galvani-Volta: A controversy that led to the discovery of useful electricity.* Norwalk, CT: Burndy Library, 1952.

Garceau, L. and H. Davis (1935). An ink-writing electroencephalogram. *Archives of Neurology and Psychiatry* 34:1292–1294.

Gazdar, G., E. Klein, G. Pullum, and I. Sag (1985). *Generalized phrase structure grammar*. Cambridge, MA: Harvard University Press.

Geddes, L. A. (2001). Contributions of the vacuum tube to early electrophysiological research. *IEEE Engineering in Medicine and Biology* pages 118–126.

Gibbs, F. A., H. Davis, and W. G. Lennox (1935). The electroencephalogram in epilepsy as a condition of impaired consciousness. *Archives of Neurology and Psychiatry* 34(6):1135–1148.

Giere, R. N. (1973). Objective single-case probabilities and the foundations of statistics. In P. Suppes, L. Henkin, G. C. Moisil, and A. Joja, eds., *Logic, methodology and philosophy of science*, vol. 4, pages 467–483. Amsterdam: North-Holland Publishing Co.

Gödel, K. (1949). A remark about the relationship between relativity theory and idealistic philosophy. In P. A. Schilpp, ed., *Albert Einstein: Philosopher-scientist*, Library of Living Philosophers, Vol. 7, pages 555–562. Chicago: Open Court Publishing.

Gogel, W. C. (1956*a*). Relative visual direction as a factor in relative distance perceptions. *Psychological Monographs* 70(11). Whole No. 418.

Gogel, W. C. (1956*b*). The tendency to see objects as equidistant and its inverse relation to lateral separation. *Psychological Monographs* 70(4). Whole No. 411.

Gogel, W. C. (1963). The visual perception of size and distance. *Vision Research* 3:101–120.

Gogel, W. C. (1964*a*). The perception of depth from binocular disparity. *Journal of Experimental Psychology* 67:379–386.

Gogel, W. C. (1964*b*). Size cue to visually perceived distance. *Psychological Bulletin* 62:217–235.

Gogel, W. C. (1965). Equidistance tendency and its consequences. *Psychological Bulletin* 64:153–163.

Goldblatt, R. (1987). *Orthogonality and space-time geometry*. New York: Springer.

Goldreich, O., S. Goldwasser, and S. Micali (1986). How to construct random functions. *Journal of the ACM* 33(4):792–807.

Good, I. J. (1962). Subjective probability as the measure of a non-measurable set. In E. Nagel, P. Suppes, and A. Tarski, eds., *Logic, methodology and philosophy of science: Proceedings of the 1960 International Congress*, pages 319–329. Stanford: Stanford University Press.

Good, I. J. (1983). *Good thinking: The foundations of probability and its applications*. Minneapolis: University of Minnesota Press.

Green, G. (1958). *An essay on the application of mathematical analysis to the theories of electricity and magnetism*. Göteborg, Sweden: Wezäta-Melins Aktiebolag. Facsimile reprint of the 1828 ed., printed for the author, by T. Wheelhouse, Nottingham.

Greenberger, D. M., M. A. Horne, A. Shimony, and A. Zeilinger (1990). Bell's theorem without inequalities. *Amer. J. Phys.* 58:1131–1143.

Greenberger, D. M., M. A. Horne, and A. Zeilinger (1989). Going beyond Bell's theorem. In M. Kafatos, ed., *Bell's theorem, quantum theory, and conceptions of the universe*. Dordrecht: Kluwer Academic Press.

Greenspan, D. G. (1973). *Discrete models*. Reading, MA: Addison–Wesley.

Greibach, F. A. (1965). A new normal form theorem for context-free phrase-structure grammars. *Journal of ACM* 12:42–52.

Griffith, R. M. (1949). Odds adjustment by American horse-race bettors. *American Journal of Psychology* 62:290–94.

Grünbaum, A. (1963). *Philosophical problems of space and time*. New York: Knopf.

Hacking, I. (1975). *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge, England: Cambridge University Press.

Hahn, R. (1967). Laplace's first formulation of scientific determinism in 1773. *Actes du XIe Congrès International d'Histoire des Sciences 1965* 2:167–171.

Hamel, G. (1958). Die Axiome der Mechanik. In *Handbuch der Physik*, vol. 5, pages 1–42. Berlin: Springer.

Hardy, L. H., G. Rand, M. C. Rittler, A. A. Blank, and P. Boeder (1953). *The geometry of binocular space perception*. Elizabeth, NJ: Schiller.

Harré, R. (1999). Models and type-hierarchies: Cognitive foundations for iconic thinking. In R. C. Paton and I. Neilson, eds., *Visual representations and interpretations*, pages 97–111. London: Springer.

Harrison, M. A. (1965). *Introduction to switching and automata theory*. New York: McGraw–Hill.

Hastie, T., R. Tibshirani, and J. Friedman (2001). *The elements of statistical learning*. New York: Springer.

Helmholtz, H. (1868). Über die Tatsachen, die der Geometrie zugrunde liegen. *Wissenschaftliche Abhandlungen* 2:618–637.

Henkin, L., P. Suppes, and A. Tarski, eds. (1959). *The axiomatic method with special reference to geometry and physics*. Amsterdam: North-Holland Publishing Co. Proceedings of an international symposium held at the University of California, Berkeley, December 16, 1957–January 4, 1958.

Hermes, H. (1938). Eine Axiomatisierung des allgemeinen Mechanik. *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften, Neue Folge* 3.

Hilbert, D. (1897/1930). *Grundlagen der Geometrie*. Leipzig and Berlin: B. G. Teubner, seventh edn. First published in 1897.

Hilbert, D. (1899). Letter to Frege. In Frege (1980), pages 39–40.

Hintikka, J. (1966). A two-dimensional continuum of inductive logic. In Hintikka and Suppes (1966), pages 113–132.

Hintikka, J. and P. Suppes, eds. (1966). *Aspects of inductive logic*. Amsterdam: North-Holland Publishing Co.

Holland, P. W. and P. R. Rosenbaum (1986). Conditional association and unidimensionality in monotone latent trait models. *Annals of Statistics* 14:1523–1543.

Hopcroft, J. E. and J. D. Ullman (1969). *Formal languages and their relation to automata*. Reading, MA: Addison–Wesley.

Hopf, E. (1934). On causality, statistics and probability. *Journal of Mathematics and Physics* 17:51–102.

Huddleston, R. and G. Pullum (2002). *The Cambridge grammar of the English language*. Cambridge, England: Cambridge University Press.

Hudgin, R. H. (1973). Coordinate-free relativity. In P. Suppes, ed., *Space, time and geometry*, pages 366–382. Dordrecht: Reidel.

Hume, D. (1739/1951). *A treatise of human nature*. London: John Noon. Quotations taken from L. A. Selby-Bigge's edition, Oxford University Press, London, 1951.

Humphreys, P. (1985). Why propensities cannot be probabilities. *Philosophical Review* 94:557–570.

Huygens, C. (1657). *De ratiociniis in ludo aleae*. Ludg. Batav., Ex officinia J. Elsevirii.

Huygens, C. (1673/1986). *The pendulum clock, or geometrical demonstration of pendula as applied to clocks*. Ames, IA: Iowa State University Press. First published in 1673. Translated by R. J. Rockwell.

Indow, T. (1967). Two interpretations of binocular visual space: Hyperbolic and Euclidean. *Annals of the Japan Association for Philosophy of Science* 3:51–64.

Indow, T. (1968). Multidimensional mapping of visual space with real and simulated stars. *Perception and Psychophysics* 3:45–53.

Indow, T. (1974*a*). Applications of multidimensional scaling in perception. In E. C. Carterette, ed., *Handbook of perception: Vol. 2. Psychophysical judgment and measurement*, pages 493–531. New York: Academic Press.

Indow, T. (1974*b*). On geometry of frameless binocular perceptual space. *Psychologia* 17:50–63.

Indow, T. (1975). An application of MDS to study of binocular visual space. In *Theory, methods and applications of multidimensional scaling and related techniques*. University of California, San Diego. U.S.-Japan Seminar (sponsored by the National Science Foundation and Japan Society for the Promotion of Science).

Indow, T. (1979). Alleys in visual space. *Journal of Mathematical Psychology* 19:221–258.

Indow, T. (1982). An approach to geometry of visual space with no a priori mapping functions: Multidimensional mapping according to Riemannian metrics. *Journal of Mathematical Psychology* 26:204–236.

Indow, T., E. Inoue, and K. Matsushima (1962*a*). An experimental study of the Luneburg theory of binocular space perception (1): The 3- and 4-point experiments. *Japanese Psychological Research* 4:6–16.

Indow, T., E. Inoue, and K. Matsushima (1962*b*). An experimental study of the Luneburg theory of binocular space perception (2): The alley experiments. *Japanese Psychological Research* 4:17–24.

Indow, T., E. Inoue, and K. Matsushima (1963). An experimental study of the Luneburg theory of binocular space perception (3): The experiments in a spacious field. *Japanese Psychological Research* 5:10–27.

James, W. (1890/1950). *Principles of psychology*. New York: Dover. First published in 1890.

Jamison, D., D. Lhamon, and P. Suppes (1970). Learning and the structure of information. In J. Hintikka and P. Suppes, eds., *Information and inference*, pages 197–259. Dordrecht: Reidel.

Jammer, M. (1961). *Concepts of mass in classical and modern physics*. Cambridge, MA: Harvard University Press.

Jasper, H. H. and L. Carmichael (1935). Electrical potentials from the intact human brain (special article). *Science* 81:51–53.

Jaynes, E. T. (1978). Where do we stand on maximum entropy? In Rosenkrantz (1983), pages 211–314.

Jaynes, E. T. (1979). Concentration of distributions at entropy maxima. In Rosenkrantz (1983), pages 315–336.

Jeffrey, R. C. (1965). *The logic of decision*. New York: McGraw-Hill.

Jeffreys, H. (1948). *Theory of probability*. Oxford: Clarendon Press, 2nd edn. First published in 1939.

Jeffreys, H. (1957). *Scientific inference*. Cambridge, England: Cambridge University Press, 2nd edn.

Jeffreys, H. and D. Wrinch (1919). On certain aspects of the theory of probability. *Philosophical Magazine* 38:715–731.

Jones, A. (2000). Pappus' notes to Euclid's Optics. In P. Suppes, J. M. Moravcsik, and H. Mendell, eds., *Ancient and mediaeval traditions in the exact sciences*, pages 49–58. Stanford, CA: CSLI Publications.

Jones, O. (1867). *Examples of Chinese ornament selected from objects in the South Kensington Museum and other collections*. London: S. & T. Gilbert.

Joyce, J. (1934). *Ulysses*. New York: Random House.

Kac, M. (1959). *Probability and related topics in the physical sciences*. New York: Interscience.

Kandel, E. R. (1985). Cellular mechanisms of learning and the biological basis of individuality. In E. R. Kandel and J. H. Schwartz, eds., *Principles of neural science*. Amsterdam: Elsevier, 2nd edn.

Kant, I. (1781/1997). *Critique of pure reason*. New York: Cambridge University Press. First published in 1781. Translated by P. Guyer and A. W. Wood.

Kant, I. (1786/1970). *Metaphysical foundations of natural science*. New York: Bobbs-Merrill Company, Inc. First published in 1786. Translated, with introduction and essay, by J. Ellington.

Kaplan, R. M. and J. Bresnan (1982). Lexical-functional grammar: A formal system for grammatical representation. In J. Bresnan, ed., *The Mental representation of grammatical relations*, pages 173–281. Cambridge, MA: MIT Press.

Kaufmann, S. A. (1982). The crystallization and selection of dynamical order in the evolution of metazoan gene regulation. In H. Haken, ed., *Evolution of order and chaos in physics, chemistry, and biology*, page 28. New York: Springer.

Keller, J. B. (1986). The probability of heads. *American Mathematical Monthly* 93:191–197.

Kelly, F. P. (1979). *Reversibility and stochastic networks*. New York: Wiley.

Kelvin (1846). On the mathematical theory of electricity in equilibrium. *Cambridge and Dublin Mathematical Journal* 1:75–96. (Lord Kelvin, also W. Thomson).

Keynes, J. M. (1921). *A Treatise on probability*. London: Macmillan.

Khinchin, A. I. (1949). *Statistical mechanics*. New York: Dover.

Kieras, D. E. (1976). Finite automata and S-R models. *Journal of Mathemathical Psychology* 13:127–147.

Kleene, S. C. (1936). Lambda-definability and recursiveness. *Duke Mathematical Journal* 2:340–353.

Kleene, S. C. (1956). Representation of events in nerve nets and finite automata. In C. E. Shannon and J. McCarthy, eds., *Automata studies*, pages 3–41. Princeton, NJ: Princeton University Press.

Klein, F. (1893). A comparative review of recent researches in geometry. *Bulletin of the New York Mathematical Society* 2:215–249.

Knorr, W. R. (1975). *The evolution of the Euclidean elements*. Dordrecht: Reidel.

Knorr, W. R. (1986). *The ancient tradition of geometric problems*. Boston, MA: Birkhäuser.

Kolmogorov, A. N. (1933/1950). *Foundations of the theory of probability*. New York: Chelsea Publishing Company. Translated by N. Morrison. First published in German in 1933. The transcription of Russian into English for the last syllable of Kolmogorov's name was generally changed from *-off* to *-ov* around 1970.

Kolmogorov, A. N. (1959). Entropy per unit time as a metric invariant of automorphisms (Russian). *Doklady Akademii Nauk SSSR* 124:754–755.

Kolmogorov, A. N. (1963). On tables of random numbers. *Sankhyā: The Indian Journal of Statistics: Series A* 25:369–376.

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problemy Peredači Informacii* 1:4–7.

Kolmogorov, A. N. (1968). Logical basis for information theory and probability theory. *IEEE Trans. Information Theory* IT-14:662–664.

Koopman, B. O. (1940*a*). The axioms and algebra of intuitive probability. *Annals of Mathematics* 41(2):269–292.

Koopman, B. O. (1940*b*). The bases of probability. *Bulletin of the American Mathematical Society* 46:763–774. Reprinted in Kyburg and Smokler (1964), 159–172.

Koopman, B. O. (1941). Intuitive probabilities and sequences. *Annals of Mathematics* 42(1):169–187.

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception & Psychophysics* 14:90–94.

Kosslyn, S. M. (1975). Information representation in visual images. *Cognitive Psychology* 7:341–370.

Kosslyn, S. M. (1976). Can imagery be distinguished from other forms of internal representation? Evidence from studies of information retrieval times. *Memory & Cognition* 14:93–111.

Kosslyn, S. M. (1978). Measuring the visual angle of the mind's eye. *Cognitive Psychology* 10:356–389.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M., B. J. Reiser, M. J. Farah, and F. S. L. (1983). Generating visual images: Units and relation. *Journal of Experimental Psychology, general* 112:278–303.

Kraft, C. H., J. W. Pratt, and A. Seidenberg (1959). Intuitive probability on finite sets. *The Annals of Mathematical Statistics* 30:408–419.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971). *Foundations of measurement Vol. I: Additive and polynomial representations*. New York: Academic Press.

Krauss, P. H. (1969). Representation of symmetric probability models. *J. Symbolic Logic* 34:183–193.

Kuratowski, C. (1921). Sur la notion de l'ordre dans la theorie des ensembles. *Fundamenta Mathematicae* 2:161–171.

Kyburg Jr., H. E. (1974). *The logical foundations of statistical inference*. Dordrecht: Reidel.

Kyburg Jr., H. E. and H. E. Smokler, eds. (1964). *Studies in subjective probability*. New York: Wiley.

Lagrange, J. L. (1873). *Oeuvres de Lagrange*, vol. 6. Paris: Gauthier-Villars. Edited by J. A. Serret.

Lamb, H. (1919). The kinematics of the eye. *Philosophical Magazine* 38:685–695.

Lamperti, J. and P. Suppes (1959). Chains of infinite order and their application to learning theory. *Pacific Journal of Mathematics* 9:739–754.

Langley, P. and J. G. Carbonell (1987). Language acquisition and machine learning. In B. MacWhinney, ed., *Mechanisms of language acquisition*, pages 115–155. Hillsdale, NJ: Lawrence Erlbaum.

Laplace, P. S. (1773/1776). Recherches, 1 sur l'intégration des équations différentielles aux différences finies, et sur leur usage dans la théorie des hasards. 2, sur le principe de la gravitation universelle, et sur les inégalités séculaires des planètes qui en dépendent. *Mémoires de mathématique et de physique présentés á l'Académie Royale des Sciences par divers Sçavans, année 1773 (often cited as Savants étrangers)* 7:37–232.

Laplace, P. S. (1774*a*). Memoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards. *Mémoires de mathématique et de physique présentés á l'Académie Royale des Sciences par divers Sçavans (often cited as Savants étrangers)* 6:353–371.

Laplace, P. S. (1774*b*). Memoire des la probabilité des causes par les événements. *Mémoires de mathématique et de physique présentés á l'Académie Royale des Sciences par divers Sçavans (often cited as Savants étrangers)* 6:621–656.

Laplace, P. S. (1799/1966). *Mécanique céleste*. Bronx, N.Y.: Chelsea Pub. Co. Vols. 1-4 is a reprint of the English translation by Nathaniel Bowditch of the edition published in Boston, 1829-1839.

Laplace, P. S. (1812). *Théorie analytique des probabilités*. Paris.

Laplace, P. S. (1951). *A philosophical essay on probabilities*. New York: Dover, sixth edn. Introduction of his *Théorie analytique des probabilités* (Laplace 1812), translated by F. W. Truscott and F. L. Emory.

Latzer, R. W. (1972). Nondirected light signals and the structure of time. *Synthese* 24:236–280.

Leibniz, G. W. (1692/1966). Essay de dynamique. In G. H. Pertz, ed., *Gesammelte Werke*. Hildesheim, Germany: George Olms. First published in 1692.

Leibniz, G. W. (1695). Specimen dynamicum. *Acta Eruditorum* 14:145–157. Translation by R. Francks and R.S. Woolhouse in G.W. Leibniz, *Philosophical texts*, Oxford University Press, New York, 1998.

Levelt, W. J. M. (1982). Cognitive styles in the use of spatial direction terms. In R. J. Jarvella and W. Klein, eds., *Speech, place, and action*, pages 251–268. New York: Wiley.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. Van Doorn, W. A. van de Grand, and K. J., eds., *Limits in perception*, pages 323–358. Utrecht: VNU Press.

Levi, I. (1980). *The enterprise of knowledge*. Cambridge, MA: MIT Press.

Lewes, G. H. (1847). The reality of Jane Eyre. *Fraser's Magazine* pages 690–693. Reprinted in "Jane Eyre", by Charlotte Brönte (1971), Richard J. Dunn's edition, p. 448.

Li, M. and P. B. M. Vitanyi (1993). *An introduction to Kolmogorov complexity and its applications*. New York: Springer.

Lindberg, D. C., ed. (1970). *John Pecham and the science of optics: Perspectiva communis*. Madison, WI: University of Wisconsin Press.

Lindenbaum, A. and A. Tarski (1934-35/1983). On the limitations of the means of expression of deductive theories. In A. Tarski, ed., *Logic, semantics, metamathematics: Papers from 1923 to 1938*, pages 384–392. Indianapolis, IN: Hackett Pub. Co. First published in 1934-35 as Uber die Beschränktheit der Ausdrucksmittel deduktiven Theorien, in *Ergebnisse eines mathematischen Kolloquiums*, **7**, pp. 15-22.

Lindley, D. V. (1971). *Bayesian statistics: A review*. Philadelphia: Soc. For Industrial and Applied Mathematics.

Lindsay, R. B. and H. Margenau (1936). *Foundations of physics*. New York: Wiley.

Loève, M. (1978). *Probability theory II*. New York: Springer, 4th edn.

Lorentz, H. A. (1892). The relative motion of the earth and the ether. *Versl. K. Ak. W. Amsterdam* 1:74. Also in H. A. Lorentz *Collected papers*, Vol. 4 p. 219–223, Nijhoff, the Hague, 1937.

Lorentz, H. A. (1895). *Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpern*. Leiden: E. J. Brill. Also in H.A. Lorentz *Collected papers*, Vol. 5 p. 1–138, Nijhoff, the Hague, 1937.

Lorentz, H. A., A. Einstein, H. Minkowski, and H. Weyl (1923). *The principle of relativity. A collection of original memoirs on the special and general theory of relativity*. New York: Dover. With notes by A. Sommerfeld. Translated by W. Perrett and G. B. Jeffery.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Luce, R. D. (1968). On the numerical representation of qualitative conditional probability. *Annals of Mathematical Statistics* 39:481–491.

Luce, R. D. (2000). *Utility of gains and losses: Measurement, theoretical and experimental approaches*. Mahwah, NJ: Erlbaum Associates.

Luce, R. D., D. H. Krantz, P. Suppes, and A. Tversky (1990). *Foundations of measurement Vol. III: Representation, axiomatization, and invariance*. San Diego, CA: Academic Press.

Luce, R. D. and P. Suppes (1965). Preference, utility and subjective probability. In R. D. Luce, R. R. Bush, and E. H. Galanter, eds., *Handbook of mathematical psychology*, vol. 3, pages 249–410. New York: Wiley.

Lucretius (1940). De rerum natura. In Oates (1940), pages 69–219.

Luneburg, R. K. (1947). *Mathematical analysis of binocular vision*. Princeton, NJ: Princeton University Press.

Luneburg, R. K. (1948). Metric methods in binocular visual perception. In *Studies and essays presented to R. Courant on his 60th birthday*, pages 215–240. New York: Wiley(Interscience).

Luneburg, R. K. (1950). The metric of binocular visual space. *Journal of the Optical Society of America* 40:627–642.

Mach, E. (1883/1942). *The science of mechanics, a critical and historical account of its development*. Chicago: Open Court Publishing, 5th edn. First published in 1883. Translated from the German by Thomas J. McCormack, containing additions and alterations up to the ninth (final) edition, with two hundred and fifty cuts and illustrations.

MacLane, S. and G. D. Birkhoff (1967). *Algebra*. New York: Macmillan.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hilbon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler (1984). *The forecasting accuracy of major time series methods*. Chichester: Wiley.

Markov, A. A. (1951). The theory of algorithms (Russian). *Trudy Matematicheskogo Instituta imeni V. A. Steklova* 38:176–189.

Martin-Löf, P. (1966). The definition of random sequences. *Information and Control* 9:606–619.

Martin-Löf, P. (1969). Literature on Von Mises' kollektivs revisited. *Theoria* 35:12–37.

Matsushima, K. and H. Noguchi (1967). Multidimensional representation of binocular visual space. *Japanese Psychological Research* 9:85–94.

Matteucci, C. (1844). *Traité des phenomènes electrophysiologiques des animaux*. Paris: Fortin Masson.

Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. *Philosophical Magazine* Reprinted in Maxwell (1890), Vol. 1, 377–409.

Maxwell, J. C. (1861-1862). On physical lines of force. *Philosophical Magazine, Series 4* 21, 23:161–175, 281–291 and 338–348, 12–24, 85–95. Reprinted in Maxwell (1890), Vol. 1, 451–513.

Maxwell, J. C. (1867). On the dynamical theory of gases. *Philosophical Transactions* 157:49–88. Reprinted in Maxwell (1890), Vol. 2, 26–78.

Maxwell, J. C. (1877). *Matter and motion*. London. Reprinted by Macmillian, London, in 1920.

Maxwell, J. C. (1890). *The scientific papers of James Clerk Maxwell*. Cambridge, England: Cambridge University Press. Edited by W. D. Niven. Reprinted by Dover, New York, in 1965.

Mayo, D. G. (1996). *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.

McGlothlin, W. H. (1956). Stability of choices among uncertain alternatives. *American Journal of Psychology* 69:604–615.

McKinsey, J. C. C., A. C. Sugar, and P. Suppes (1953). Axiomatic foundations of classical particle mechanics. *Journal of Rational Mechanics and Analysis* 2:253–272.

McKinsey, J. C. C. and P. Suppes (1953). Transformations of systems of classical particle mechanics. *Journal of Rational Mechanics and Analysis* 2:273–289.

Mehlberg, H. (1935). Essai sur la théorie causale du temps. I. *Studia Philosophica* 1:119–260.

Mehlberg, H. (1937). Essai sur la théorie causale du temps. II. *Studia Philosophica* 2:111–231.

Mellor, D. H. (1971). *The matter of chance*. Cambridge, England: Cambridge University Press.

Mermin, N. D. (1990*a*). Quantum mysteries revisited. *American Journal of Physics* 58(8):731–734.

Mermin, N. D. (1990*b*). Extreme quantum entanglement in a superposition of macroscopically distinct states. *Physical Review Letters* 65(15):1838–1840.

Mermin, N. D. (1993). Hidden variables and the two theorems of John Bell. *Reviews of Modern Physics* 65(3):803–815.

Michelson, A. A. and E. W. Morley (1887). On the relative motion of the earth and the luminiferous ether. *American Journal of Science* 34:333–336.

Millenson, J. R. (1967). An isomorphism between stimulus-response notation and information processing flow diagrams. *The Psychological Record* 17:305–319.

Miller, G. A. and N. Chomsky (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, and E. Galanter, eds., *Handbook of Mathematical Psychology*, vol. 2. New York: Wiley.

Miller, G. A., E. Galanter, and K. H. Pribram (1960). *Plans and the structure of behavior*. New York: Holt.

Milne, E. A. (1948). *Kinematic relativity*. London: Oxford University Press.

Montgomery, R. (2001). A new solution to the three-body problem. *Notices of the American Mathematical Society* 48:471–481.

Moody, E. A. and M. Clagett, eds. (1952). *The medieval science of weights (Scientia de ponderibus)*. Madison, WI: The University of Wisconsin Press.

Moser, J. (1973). *Stable and random motions in dynamical systems with special emphasis on celestial mechanics*. Herman Weyl Lectures, the Institute for Advanced Study. Princeton, NJ: Princeton University Press.

Mosteller, F. and D. L. Wallace (1964/1984). *Applied Bayesian and classical inference: The case of the federalist papers*. Springer Series in Statistics. New York: Springer.

Moulines, C. U. (1975). A logical reconstruction of classical equilibrium thermodynamics. *Erkenntnis* 9:101–130.

Moulines, C. U. (1976). Approximate application of empirical theories: A general explication. *Erkenntnis* 10:201–227.

Moulines, C. U. and J. D. Sneed (1979). Suppes' philosophy of physics. In R. J. Bogdan, ed., *Patrick Suppes*, pages 59–91. Dordrecht: Reidel.

Mundy, B. (1986*a*). Optical axiomatization of Minkowski space-time geometry. *Philosophy of Science* 37:1–30.

Mundy, B. (1986*b*). The physical content of Minkowski geometry. *British Journal for the Philosophy of Science* 37:25–54.

Murray, N. and M. Holman (2001). The role of chaotic resonances in the solar system. *Nature* 410:773–779.

Nagel, E. (1939*a*). The formation of modern conceptions of formal logic in the development of geometry. *Osiris* 7:142–224. Reprinted in E. Nagel *Teleology revisited*, New York, Columbia University Press, 1979, pp. 195–259.

Nagel, E. (1939*b*). Principles of the theory of probability. In *International Encyclopedia of Unified Science*, vol. 1 (6). Chicago, IL: University of Chicago Press.

Nagel, E. (1949). The meaning of reduction in the natural sciences. In R. C. Stauffer, ed., *Science and civilization*, pages 99–135. Madison, WI: University of Wisconsin Press.

Neimark, E. D. and W. K. Estes, eds. (1967). *Stimulus sampling theory*. Holden-Day series in Psychology. San Francisco: Holden-Day.

Nelson, E. (1987). *Radically elementary probability theory*. Princeton, NJ: Princeton University Press.

Neugebauer, O. (1957). *The exact sciences in antiquity*. Providence, RI: Brown University Press, 2nd edn.

Newton, I. (1687/1946). *Principia*. Berkeley, CA: University of California Press. First published in 1687. Translation by F. Cajori.

Newton, I. (1704/1931). *Opticks*. London: Bell. Reprinted from the 4th edition. First published in 1704.

Neyman, J. (1971). Foundations of behavioristic statistics. In V. Godambe and D. Sprott, eds., *Foundations of statistical inference*, pages 1–19. Toronto: Holt, Rinehart and Winston of Canada.

Nishikawa, Y. (1967). Euclidean interpretation of binocular visual space. *Japanese Psychological Research* 9:191–198.

Noether, E. (1918). Invariante Variationsprobleme. *Nachr. Konig. Gesell. Wissen. Göttingen, Math.-Phys. Klasse 2* pages 235–257. Reprinted in Noether *Collected papers*, Springer, New York 1983.

Noll, W. (1955). Die Herleitung der Grundgleichungen der Thermomechanik der Kontinua aus der statistischen Mechanik. *Journal of Rational Mechanics and Analysis* 4:627–646.

Noll, W. (1974). *The foundations of mechanics and thermodynamics: Selected papers by W. Noll*. New York: Springer. With a preface by C. Truesdell.

Norman, M. F. (1972). *Markov processes and learning models*. New York: Academic Press.

Oates, W. J., ed. (1940). *The Stoic and Epicurean philosophers: The complete extant writings of Epicurus, Epictetus, Lucretius, Marcus Aurelius*. New York: Random House, by arrangement with Oxford University Press.

Oppenheim, A. V. and R. W. Schafer (1975). *Digital signal processing*. Englewood Cliffs, NJ: Prentice-Hall.

Ornstein, D. S. (1970). Bernoulli shifts with the same entropy are isomorphic. *Advances in Mathematics* 4:337–352.

Ornstein, D. S. and B. Weiss (1991). Statistical properties of chaotic systems. *Bull. Am. Math. Soc.(New Series)* 24:11–116.

Padoa, A. (1902). Un nouveau système irréductible de postulats pour l'algèbre. *C. R. Deuxième Congrès International des Mathématiciens, Paris* pages 249–256.

Pais, A. (1982). *'Subtle is the Lord ...' The science and life of Albert Einstein*. New York: Oxford University Press.

Palladio, A. (1570/1965). *The four books of architecture*. New York: Dover. Unabridged and unaltered republication of the work first published by Isaac Ware in 1738.

Pappus (1876-1878). *Pappi Alexandrini collectionis quae supersunt*. Berlin: Weidman. Edited by F. Hultsch.

Pasch, M. (1882). *Vorlesungen über Neuere Geometrie*. Leipzig: Verlag von Julius Springer.

Peacocke, C. (1983). *Sense and content : Experience, thought, and their relations*. London: Oxford University Press.

Pecham, J. (1970). Perspectiva communis. In Lindberg (1970).

Peres, Y. (1992). Finite violation of a Bell inequality for arbitrarily large spin. *Physical Review A* 46:4413–4414.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Plato (1892). Meno. In *The dialogues of Plato*. London: Macmillian. Translated by B. Jowett. Reprinted in 1937, Random House, New York.

Poincaré, H. C. (1898). The measuring of time. *Revue de Métaphysique et de Morale* 6:1–13.

Poincaré, H. C. (1905). Sur la dynamique de l'électron. *Comptes Rendus de l'Académie des Sciences Paris* 140:1504. Reprinted in Oeuvre de Henry Poincaré, Vol. 9, p. 489, Paris, Gauthier-Villars, 1954.

Poincaré, H. C. (1906). Sur la dynamique de l'électron. *Rendiconti del Circolo Matematico di Palermo* 21:129–175. Reprinted in Oeuvre de Henry Poincaré, Vol. 9, p. 494, Paris, Gauthier-Villars, 1954.

Poincaré, H. C. (1910). *Sechs Vorträge aus der Reinen Mathematik und Mathematishen Physic*. Leipzig: Teubner.

Poincaré, H. C. (1912). *Calcul des probabilités*. Paris: Gauthier-Villars, 2nd edn.

Poisson, S. D. (1833). *Traité de mécanique*. Paris: Bachelier, 2nd edn.

Poisson, S. D. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Paris.

Poncelet, J. V. (1865). *Traité des proprietés projectives de figures*. Paris: Gauthier-Villars, 2nd edn. First published in 1822.

Popper, K. R. (1957). The propensity interpretation of the calculus of probability and the quantum theory. In S. Körner, ed., *Observation and interpretation in the philosophy of physics*. London: Butterworth.

Popper, K. R. (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science* 10:26–42.

Popper, K. R. (1974). Replies to my critics. In P. A. Schilpp, ed., *The philosophy of Karl Popper*, vol. 2 of *Library of Living Philosophers, Vol. 14*. Chicago: Open Court Publishing.

Portugal, F. H. and J. S. Cohen (1977). *A century of DNA, a history of the discovery of the structure and function of the genetic substance*. Cambridge, MA: MIT Press.

Post, E. L. (1936). Finite combinatory processes. Formulation I. *Journal of Symbolic Logic* 1:103–5.

Preston, M. G. and P. Baratta (1948). An experimental study of the auction value of an uncertain outcome. *American Journal of Psychology* 61:183–193.

Ptolemy, C. (1984). *The almagest*. New York: Springer. Translated and annotated by G. J. Toomer. Written about 150 A.D.

Rabin, M. O. (1963). Probabilistic automata. *Information and Control* 6:230–245.

Rabin, M. O. and D. Scott (1959). Finite automata and their decision problems. *IBM Journal of Research and Development* 3:114–125. Reprinted in E. F. Moore (Ed.), Sequential machines. Reading MA, Wesley, 1964, 98-114.

Ramsey, F. P. (1931). Truth and probability. In R. B. Braithwaite, ed., *The foundations of mathematics and other logical essays*, pages 156–198. London: Kegan Paul, Trench, Trubner & Co.

Redi, F. (1686). *Esperienze intorno a diverse cose naturali, e particolarmente a quelle, che ci son portate dall'Indie*. Florence: Piero Matini. First published in 1671.

Reichenbach, H. (1924). *Axiomatik der relativistichen Raum-Zeit-Lehre*. Braunschweig: Vieweg and Sons.

Reichenbach, H. (1932). Axiomatik der wahrscheinlichkeitsrechnung. *Math. Zs.* 34:568–619.

Reid, T. (1764/1967). Inquiry into the human mind. In *Philosophical works*, vol. 1. Hildesheim, Germany: George Olms. First published in 1764.

Riemann, B. (1866–1867). Über die Hypothesen, welche der Geometrie zu Grunde liegen. *Gesellschaft der Wissenschaften zu Göttingen: Abhandlungen* 13:133–152. Habilitationsschrift, Göttingen, 1854. Translation: "On the hypotheses which lie at the basis of geometry", Nature VIII, 14-17, 36f (May 1873).

Robb, A. A. (1911). *Optical geometry of motion: A new view of the theory of relativity*. Cambridge, England: Heffer.

Robb, A. A. (1914). *A theory of time and space*. New York: Cambridge University Press.

Robb, A. A. (1921). *The absolute relations of time and space*. New York: Cambridge University Press.

Robb, A. A. (1928). On the connexion of a certain identity with the extension of conical order to n dimensions. *Proceedings of the Cambridge Philosophical Society* 24:357–374.

Robb, A. A. (1930). On a symmetrical analysis of conical order and its relation to time-space. *Proceedings of the Royal Society of London, Series A* 129:549–579.

Robb, A. A. (1936). *Geometry of time and space*. New York: Cambridge University Press.

Roberts, F. S. and P. Suppes (1967). Some problems in the geometry of visual perception. *Synthese* 17:173–201.

Rogers Jr., H. (1967). *Theory of recursive functions and effective computability*. New York: McGraw–Hill.

Rollins, M. (1989). *Mental imagery: On the limits of cognitive science*. New Haven: Yale University Press.

Rosenkrantz, R. D. (1977). *Inference, method and decision*. Dordrecht: Reidel.

Rosenkrantz, R. D., ed. (1983). *E. T. Jaynes: Papers on probability, statistics and statistical physics*. Dordrecht: Reidel.

Rouanet, H., J. M. Bernard, M. C. Bert, B. Lecoutre, M. P. Lecoutre, and B. Le Roux (1998). *New ways in statistical methodology: From significance tests to Bayesian inference*. European University Studies, Series VI-Psychology, Vol. 618. Berne: Peter Lang.

Rubin, G. S. and K. Turano (1992). Reading without saccadic eye movements. *Vision Research* 32(5):895–902.

Rubin, H. and P. Suppes (1954). Transformations of systems of relativistic particle mechanics. *Pacific Journal of Mathematics* 4:563–601.

Rubin, H. and P. Suppes (1955). A note on two-place predicates and fitting sequences of measure functions. *Journal of Symbolic Logic* 20:121–122.

Ruelle, D. (1969). *Statistical mechanics: Rigorous results*. New York: W. A. Benjamin.

Ruelle, D. (1978). *Thermodynamic formalism*. Reading, MA: Addison–Wesley. Encyclopedia of mathematics and its application, Vol. 5.

Rugg, M. D. and M. G. Coles, eds. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. New York: Oxford University Press.

Salmon, W. (1979). Propensities: A discussion review. *Erkenntnis* 14:183–216.

Sambursky, S. (1956). On the possible and probable in ancient Greece. *Osiris* 12:35–48.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley. Revised edition printed in 1972, Dover, New York.

Schelling, H. (1956). Vision. *Journal of the Optical Society of America* 46:309–315.

Schilpp, P. A., ed. (1963). *The philosophy of Rudolf Carnap*. Library of Living Philosophers, Vol. 11. Chicago: Open Court Publishing.

Schutz, J. W. (1973). *Foundations of special relativity: Kinematik axioms for Minkowski space-time*. New York: Springer.

Schutz, J. W. (1979). An axiomatic system for Minkowski space-time. Tech. Rep. MPI-PAE/Astro 181, Max-Plank-Institut für Physik und Astrophysik, Munich.

Scott, D. (1964). Measurement structures and linear inequalities. *Journal of Mathematical Psychology* 1:233–247.

Scott, D. and P. Krauss (1966). Assigning probabilities to logical formulas. In Hintikka and Suppes (1966), pages 219–264.

Scott, D. and P. Suppes (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic* 23:113–128.

Shepard, R. N. (1966). Learning and recall as organization and search. *Journal of Verbal Learning and Verbal Behavior* 5:201–204.

Shepard, R. N. and L. A. Cooper (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press/Bradford.

Shepard, R. N. and J. Metzler (1971). Mental rotation of three-dimensional objects. *Science* 171:701–703.

Shepherdson, J. C. and H. E. Sturgis (1963). Computability of recursive functions. *Journal of the ACM* 10:217–255.

Shuford, E. H. (1959). A comparison of subjective probabilities for elementary and compound events. Tech. Rep. 20, The Psychometric Lab., University of North Carolina.

Siacci, F. (1878). Del moto per una linea piana. *Atti della Reale Accademia di Torino* 14:750–766.

Sierpenski, W. (1958). *Cardinal and ordinal numbers*. Warsaw: Polska Akademia Nauk, Monografie Matematyczne. Tom. 34 Panstwowe Wydawnictwo Naukowe.

Simon, H. A. (1957). *Models of man*. New York: Wiley.

Sinai, Y. G. (1959). On the concept of entropy of a dynamical system. *Dokl. Akad. Nauk. SSSR* 124:768–771.

Siskind, J. M. (1991). Dispelling myths about language bootstrapping. In *Proceedings of the AAAI Spring Symposium Workshop on Machine Learning of Natural Language and Ontology*.

Siskind, J. M. (1992). *Naive physics, event perception, lexical semantics, and language acquisition*. Ph.D. thesis, Elec. Eng. & Comp. Sci., MIT, Cambridge, MA.

Siskind, J. M. (1994). Lexical acquisition in the presence of noise and homonymy. In *Proceedings of the 12th National Conference on Artificial Intelligence, AAAI-94*, pages 760–766.

Sitnikov, K. (1960). Existence of oscillating motions for the three-body problem. *Doklady Akademii Nauk, USSR* 133(2):303–306.

Smith, C. A. B. (1961). Consistency in statistical inference and decision. *J. R. Statist. Soc., B* 23:1–25.

Smoluchowski, M. (1918). Über den Begriff des Zufalls und den Ursprung der Wahrscheinlichkeitsgesetze in der Physik. *Die Naturwissenschaften* pages 253–263.

Snapper, E. and R. J. Troyer (1971). *Metric affine geometry*. New York: Academic Press.

Sneed, J. D. (1971). *The logical structure of mathematical physics*. Dordrecht: Reidel.

Soare, R. I. (1987). *Recursively enumerable sets and degrees*. New York: Springer.

Solomonoff, R. J. (1964). A formal theory of inductive inference. *Information and Control* 7:1–22, 224–254.

Sommer, R. and P. Suppes (1996). Finite models of elementary recursive nonstandard analysis. *Notas de la Sociedad Matematica de Chile* 15:73–95.

Sommer, R. and P. Suppes (1997). Dispensing with the continuum. *Journal of Mathematical Psychology* 41:3–10.

Stegmüller, W. (1976). *The structure and dynamics of theories*. New York: Springer. Translation of *Theorienstrukturen und Theoriendynamik*, originally published as V. 2, Pt. 2 of the author's *Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie*, 1973.

Stegmüller, W. (1979). *The structuralist view of theories: A possible analogue of the Bourbaki programme in physical science*. Berlin: Springer.

Stein, W. (1930). Der Begriff des Schwerpunktes bei Archimedes. *Ouellen und Studien zur Geschichte der Mathematik, Physik und Astronomie* 1:221–224.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science* 103:677–680.

Strawson, P. F. (1966). *The bounds of sense: An essay on Kant's Critique of Pure Reason*. London: Methuen.

Suppe, F., ed. (1974). *The structure of scientific theories*. Urbana: University of Illinois Press.

Suppes, P. (1956). The role of subjective probability and utility in decision-making. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954-1955*, vol. 5, pages 61–73.

Suppes, P. (1957/1999). *Introduction to logic*. New York: Van Nostrand. Reprinted in 1999 by Dover, New York.

Suppes, P. (1959). Axioms for relativistic kinematics with or without parity. In Henkin et al. (1959), pages 291–307.

Suppes, P. (1960*a*). A comparison of the meaning and uses of models in mathematics and the empirical sciences. *Synthese* 12:287–301.

Suppes, P. (1960*b*). *Axiomatic set theory*. New York: Van Nostrand. Slightly revised edition published by Dover, New York, 1972.

Suppes, P. (1962). Models of data. In E. Nagel, P. Suppes, and A. Tarski, eds., *Logic, methodology and philosophy of science: Proceedings of the 1960 International Congress*, pages 252–261. Stanford: Stanford University Press.

Suppes, P. (1969*a*). *Studies in the methodology and foundations of science: Selected papers from 1951 to 1969*. Dordrecht: Reidel.

Suppes, P. (1969*b*). Stimulus-response theory of finite automata. *Journal of Mathematical Psychology* 6:327–355.

Suppes, P. (1969*c*). Stimulus-response theory of finite automata and TOTE hierarchies: A reply to Arbib. *Psychological Review* 76:511–514.

Suppes, P. (1970*a*). *A probabilistic theory of causality*. Amsterdam: North-Holland Publishing Co.

Suppes, P. (1970*b*). Probabilistic grammars for natural languages. *Synthese* 22:95–116.

Suppes, P. (1972). Finite equal-interval measurement structures. *Theoria* 38:45–63.

Suppes, P. (1973*a*). New foundations of objective probability: Axioms for propensities. In P. Suppes, L. Henkin, G. C. Moisil, and A. Joja, eds., *Logic, methodology, and philosophy of science IV: Proceedings of the Fourth International Congress for Logic, Methodology and Philosophy of Science, Bucharest, 1971*, pages 515–529. Amsterdam: North-Holland Publishing Co.

Suppes, P. (1973*b*). Congruence of meaning. *Proceedings and Addresses of the American Philosophical Association* 46:21–38.

Suppes, P. (1974*a*). Popper's analysis of probability in quantum mechanics. In P. A. Schilpp, ed., *The philosophy of Karl Popper*, pages 760–774. Chicago: Open Court Publishing.

Suppes, P. (1974*b*). The structure of theories and the analysis of data. In F. Suppe, ed., *The structure of scientific theories*, pages 266–283. Urbana, IL: University of Illinois Press, 2nd edn.

Suppes, P. (1974*c*). Aristotle's concept of matter and its relation to modern concepts of matter. *Synthese* 28:27–50.

Suppes, P. (1974*d*). The measurement of belief. *Journal of the Royal Statistical Society (Series B)* 36:160–191.

Suppes, P. (1975). From behaviorism to neobehaviorism. *Theory and Decision* 6:269–285.

Suppes, P. (1976). Testing theories and the foundations of statistics. In W. L. Harper and C. A. Hooker, eds., *Foundations of probability theory, statistical inference, and statistical theories of science*, vol. 2, pages 437–455. Dordrecht: Reidel.

Suppes, P. (1977*a*). Is visual space Euclidean? *Erkenntnis* 36:397–421.

Suppes, P. (1977*b*). Learning theory for probabilistic automata and register machines. In H. Spada and W. F. Kempf, eds., *Structural models of thinking and learning*, pages 57–79. Bern: Hans Huber Publisher. Proceedings of the 7th IPN-Symposium on formalized theories of thinking and learning and their implications for science instruction.

Suppes, P. (1979). The logic of clinical judgment: Bayesian and other approaches. In H. T. Engelhardt, Jr., S. F. Spicker, and B. Towers, eds., *Clinical judgment: A critical appraisal*, pages 145–159. Dordrecht: Reidel.

Suppes, P. (1980). Limitations of the axiomatic method in ancient Greek mathematical sciences. In J. Hintikka, D. Gruender, and E. Agazzi, eds., *Pisa Conference Proceedings*, vol. 1, pages 197–213. Dordrecht: Reidel.

Suppes, P. (1983). Arguments for randomizing. In P. D. Asquith and T. Nickles, eds., *PSA 1982*, pages 464–475. Lansing, MI: Philosophy of Science Association.

Suppes, P. (1984). *Probabilistic metaphysics*. Oxford: Blackwell.

Suppes, P. (1987). Propensity representations of probability. *Erkenntnis* 26:335–358.

Suppes, P. (1988). Empirical structures. In E. Scheibe, ed., *The role of experience in science, proceedings of 1986 conference of the Académie Internationale de Philosophie des Sciences (Bruxelles)*, pages 23–33. New York: Walter de Gruyter.

Suppes, P. (1989). Current directions in mathematical learning theory. In E. E. Roskam, ed., *Mathematical Psychology in Progress*, pages 3–28. New York: Springer.

Suppes, P. (1991). *Language for humans and robots*. Oxford: Blackwell.

Suppes, P. (1995). Some foundational problems in the theory of visual space. In R. D. Luce, M. D'Zmura, D. Hoffman, G. J. Iverson, and A. K. Romney, eds., *Geometric representations of perceptual phenomena: Papers in honor of Tarow Indow on his 70th birthday*, pages 37–45. Mahwah, NJ: Erlbaum Associates.

Suppes, P. (1998). Pragmatism in physics. In P. Weingartner, G. Schurz, and G. Dorn, eds., *The role of pragmatics in contemporary philosophy*, pages 236–253. Vienna: Holder, Pichler, Tempsky.

Suppes, P. (2000). Quantifier-free axioms for constructive affine plane geometry. *Synthese* 125:263–281.

Suppes, P. (2001*a*). Finitism in geometry. *Erkenntnis* 54:133–144.

Suppes, P. (2001*b*). Weak and strong reversibility of causal processes. In M. C. Galavotti, P. Suppes, and D. Costantini, eds., *Stochastic causality*, pages 203–220. Stanford, CA: CSLI Publications.

Suppes, P. and R. C. Atkinson (1960). *Markov learning models for multiperson interactions*. Stanford, CA: Stanford University Press.

Suppes, P., M. Böttner, and L. Liang (1995). Comprehension grammars generated from machine learning of natural language. *Machine Learning* 19:133–152.

Suppes, P., M. Böttner, and L. Liang (1996). Machine learning comprehension grammars for ten languages. *Computational Linguistics* 22:329–350.

Suppes, P. and R. Chuaqui (1993). A finitarily consistent free-variable positive fragment of infinitesimal analysis. *Notas de Logica Matematica* 38:1–59. Proceedings of the IX Latin American Symposium on Mathematical Logic, held at Bahia Blanca, Argentina, August 1992.

Suppes, P., A. J. de Barros, and G. Oas (1998). A collection of probabilistic hidden-variable theorems and counterexamples. In R. Pratesi and L. Ronchi, eds., *Waves, information and foundations of physics. Conference proceedings, Vol. 60*, pages 267–291. Bologna: Società Italiana Di Fisica.

Suppes, P., B. Han, J. Epelboim, and Z. L. Lu (1999*a*). Invariance between subjects of brain wave representations of language. *Proceedings of the United States National Academy of Sciences* 96:12953–12958.

Suppes, P., B. Han, J. Epelboim, and Z. L. Lu (1999*b*). Invariance of brain-wave representations of simple visual images and their names. *Proceedings of the United States National Academy of Sciences* 96:14658–14663.

Suppes, P., B. Han, and Z. L. Lu (1998). Brain-wave recognition of sentences. *Proceedings of the National Academy of Sciences* 95:15861–15866.

Suppes, P., D. H. Krantz, R. D. Luce, and A. Tversky (1989). *Foundations of measurement volume II: Geometrical, threshold, and probabilistic representations*. San Diego, CA: Academic Press.

Suppes, P. and L. Liang (1996). Probabilistic association and denotation in machine learning of natural language. In A. Gammerman, ed., *Computational learning and probabilistic reasoning*, pages 87–100. New York: Wiley.

Suppes, P. and L. Liang (1998). Concept learning rates and transfer performance of several multivariate neural network models. In C. E. Dowling, F. S. Roberts, and P. Theuns, eds., *Recent progress in mathematical psychology*, pages 227–252. Mahwah, NJ: Erlbaum Associates.

Suppes, P., L. Liang, and M. Böttner (1992). Complexity issues in robotic machine learning of natural language. In L. Lam and V. Naroditsky, eds., *Modelling complex phenomena*, pages 102–127. New York: Springer.

Suppes, P., Z. L. Lu, and B. Han (1997). Brain-wave representations of words. *Proceedings of the United States National Academy of Sciences* 94:14965–14969.

Suppes, P. and W. Rottmayer (1974). Automata. In E. C. Carterette and M. P. Friedman, eds., *Handbook of perception, Vol. 1: Historical and philosophical roots of perception*, pages 335–362. New York: Academic Press.

Suppes, P. and M. Winet (1955). An axiomatization of utility based on the action of utility differences. *Journal of Management Science* 1:259–270.

Suppes, P., D. K. Wong, M. Perreau-Guimaraes, E. T. Uy, and W. Yang (to appear). High statistical recognition rates for some persons' brain-wave representations of sentences .

Suppes, P. and M. Zanotti (1976). Necessary and sufficient conditions for existence of a unique measure strictly agreeing with a qualitative probability ordering. *Journal of Philosophical Logic* 5:431–438.

Suppes, P. and M. Zanotti (1980). A new proof of the impossibility of hidden variables using the principles of exchangeability and identity of conditional distribution. In P. Suppes, ed., *Studies in the foundations of quantum mechanics*, pages 173–191. East Lansing, MI: Philosophy of Science Association.

Suppes, P. and M. Zanotti (1981). When are probabilistic explanations possible? *Synthese* 48:191–199.

Suppes, P. and M. Zanotti (1982). Necessary and sufficient qualitative axioms for conditional probability. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 60:163–169.

Suppes, P. and M. Zanotti (1996). Mastery learning of elementary mathematics: Theory and data. In *Foundations of probability with applications: Selected papers, 1974–1995*, pages 149–188. New York: Cambridge University Press.

Suppes, P. and J. L. Zinnes (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, and E. H. Galanter, eds., *Handbook of mathematical psychology*, vol. 1, pages 3–76. New York: Wiley.

Swerdlow, N. M. (1998). *The Babylonian theory of the planets*. Princeton: Princeton University Press.

Szekeres, G. (1968). Kinematic geometry: An axiomatic system for Minkowski space-time. *Journal of Australian Mathematical Society* 8:134–160.

Szmielew, W. (1983). *From affine to Euclidean geometry: An axiomatic approach*. Warszawa, Poland and Dordrecht, Holland: PWN-Polish Scientific Publishers and D. Reidel.

Tait, W. W. (1959). A counterexample to a conjecture of Scott and Suppes. *Journal of Symbolic Logic* 24:15–16.

Tarski, A. (1935). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica* 1:261–405.

Tarski, A. (1953). *Undecidable theories*. Amsterdam: North-Holland Publishing Co.

Tarski, A. and S. Givant (1987). *A formalization of set theory without variables*. Providence, RI: American Mathematical Society.

Toda, M. (1951). Measurement of intuitive-probability by a method of game. *Japan. J. Psychol.* 22:29–40.

Toda, M. (1958). Subjective inference vs. objective inference of sequential dependencies. *Japan. Psychol. Res.* 5:1–20.

Todhunter, I. (1865/1949). *A history of the mathematical theory of probability from the time of Pascal to that of Laplace*. New York: Chelsea Publishing Company. First published in 1865.

Toth, L. F. (1964). *Regular figures*. New York: Macmillan.

Truesdell, C. (1968). *Essays in the history of mechanics*. New York: Springer.

Tulving, E. and F. I. M. Craik, eds. (2000). *The Oxford handbook of memory*. New York: Oxford University Press.

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proc. London Math. Soc. Ser. 2* 42:230–265. A correction, Vol. 43 (1936), 544–546.

Tye, M. (1991). *The imagery debate*. Representation and mind. Cambridge, MA: MIT Press/Bradford.

van der Waerden, B. L., ed. (1968). *Sources of quantum mechanics*. New York: Dover.

van Fraassen, B. C. (1980). *The scientific image*. Oxford: Clarendon Press.

van Lambalgen, M. (1987*a*). *Random sequences*. Amsterdam: Dutch Foundation for Scientific Research.

van Lambalgen, M. (1987*b*). Von Mises' definition of random sequences reconsidered. *Journal of Symbolic Logic* 52(3):725–755.

van Lambalgen, M. (1990). The axiomatization of randomness. *Journal of Symbolic Logic* 55(3):1143–1167.

van Lambalgen, M. (1992). Independence, randomness and the axiom of choice. *Journal of Symbolic Logic* 57(4):1274–1304.

van Lambalgen, M. (1996). Randomness and foundations of probability: Von Mises' axiomatisation of random sequences. In T. S. Ferguson, L. S. Shapley, and J. B. MacQueen, eds., *Statistics, probability and game theory. Papers in honor of David Blackwell*, Lecture Notes - Monograph Series, Volume 30, pages 347–367. Hayward, CA: Institute of Mathematical Statistics.

Vaught, R. (1954). Remarks on universal classes of relational systems. *Indagationes Matematicae* 16:589–591.

Veblen, O. (1904). A system of axioms for geometry. *Transaction of the American Mathematical Society* 5:343–384.

Veblen, O. and J. W. Young (1910). *Projective geometry - Vol. 1*. Boston, MA: Ginn. Reprinted by Blaisdell, New York, 1938.

Veblen, O. and J. W. Young (1918). *Projective geometry - Vol. 2*. Boston, MA: Ginn. Reprinted by Blaisdell, New York, 1938.

Ville, J. (1939). *Ètude critique de la notion de collectif*. Paris: Gauthiers-Villars.

Vitruvius (1960). *The ten books on architecture*. New York: Dover. Originally published by Harvard University Press in 1914. Translated by Morris Hicky Morgan.

Voigt, W. (1887). Über das Doppler'sche Prinzip. *Nachr. Ges. Wiss. Göttingen* 41.

Volta, A. (1793/1918). Letter to Tiberius Cavallo, 22 May 1793. In *Le opere di Alessandro Volta*, vol. 1, pages 203–208. Milan: Ulrico Hoepli.

von Mises, R. (1919). Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift* 5:52–99.

von Mises, R. (1941). On the foundations of probability and statistics. *Ann. Math. Stat.* 12:191–205, 215–216.

von Neumann, J. (1932/1955). *Mathematical foundations of quantum mechanics*. Princeton, NJ: Princeton University Press. Translated from the German edition of 1932 by R. T. Byer.

von Plato, J. (1983). The method of arbitrary functions. *The British Journal for the Philosophy of Science* 34:37–47.

Wagner, M. (1985). The metric of visual space. *Perceptions & Psychophysics* 38:483–495.

Wald, A. (1936). Sur la notion de collectif dans le calcul des probabilités. *C. R. Acad. Sci. Paris* 202:180–183.

Walker, A. G. (1948). Foundations of relativity. Parts I and II. *Proceedings of the Royal Society of Edimburgh, Section A, Mathematical and Physical Sciences* 62:319–335.

Walker, A. G. (1959). Axioms for cosmology. In Henkin et al. (1959), pages 308–321.

Wexler, K. and P. W. Culicover (1980). *Formal principles of language aquisition*. Cambridge, MA: MIT Press.

Weyl, H. (1928/1931). *The theory of groups and quantum mechanics*. London: Methuen & Co. Ltd. Translation from the second (revised) German edition of *Gruppentheorie und Quantenmechanik*, 1928.

Whittaker, E. T. (1904/1937). *A treatise on the analytical dynamics of particles and rigid bodies*. Cambridge, England: Cambridge University Press, 4th edn. First published in 1904.

Whittaker, E. T. (1910). *A history of the theories of aether and electricity*. London: Longmans, Green and Co.

Wiener, N. (1914). A simplification of the logic of relations. *Proceedings of the Cambridge Philosophical Society* 17:387–390.

Wilson, E. B. and G. N. Lewis (1912). The space-time manifold of relativity. The non-Euclidean geometry of mechanics and electromagnetics. *Proceedings of the American Academy of Arts and Sciences* 48:389–507.

Winnie, J. A. (1977). The causal theory of space-time. In J. Earman, C. Glymour, and J. Stachel, eds., *Foundations of space-time theories*, pages 134–205. Minneapolis: University of Minnesota Press.

Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. New York: Routledge. English and German.

Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media Inc.

Wussing, H. (1984). *The genesis of the abstract group concept*. Cambridge, MA: MIT Press.

Zajaczkowska, A. (1956). Experimental test of Luneburg's theory: Horopter and Alley experiments. *Journal of the Optical Society of America* 46:514–527.

Zeeman, E. C. (1964). Causality implies the Lorentz group. *Journal of Mathematical Physics* 5:490–493.

Zeeman, E. C. (1967). The topology of Minkowski space. *Topology* 6:161–170.

Zvonkin, A. K. and L. A. Levin (1970). The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematical Surveys* 25:83–124.

术 语 索 引

（后面所标页码为原书页码，即本书边码）

译 后 记

本书的翻译工作开始于 2008 年 2 月 15 日。当时,我在作者领导的研究小组进行为期一年的学术访问,主要从事科学哲学与物理哲学研究。本书前五章的初译稿是在作者身边完成的,那段时间,他每天下午都会到我的办公室转一圈,随时与我讨论翻译过程中遇到的问题。后面的三章是在回国后进行的,我的博士生赫中华女士参与了第六章和第八章的初译工作,并对这两章的许多图表作了技术性的处理。第八章是技术性很强的一章,为了保证技术用语与理解上的准确性,作者请他的博士生王瑞女士对第八章的第一节和第六节作了校译,特别是修正了某些专业术语的翻译,并在作者认可的前提下,修改了原文中某些定理的错误之处。王瑞是中国人,她曾在清华大学获得学士与硕士学位,当时,师从作者从事神经科学方面的研究。2009 年 7 月 20 日在作者的邀请与资助下,我再次有机会赴斯坦福大学进行了为期一个月的合作研究工作。在合作研究期间,解决了后面部分翻译过程中遇到的问题。

本书在翻译期间还得到我的室友德拉勒·斯特拉特曼·亚历山大(Delalle Strateman Alexander)的热情帮助,我两次访问斯坦福大学期间都与她合住在同一个公寓里。德拉勒是埃及人,1985 年在哥伦比亚大学国际与公共事务学院硕士毕业,后来,一直生活在美国,她有两个女儿,当时,大女儿在英国伦敦的

一家银行工作;小女儿在加州大学伯克利分校读本科。她年长于我,曾在生活上给予我最大的关照,并随时耐心地与我讨论翻译中的一些难以理解的英语句子。我在访问斯坦福期间,与泽千鹤子(Chiauke Izawa)教授共用一间办公室。泽千鹤子是日本人,于20世纪60年代从日本来到斯坦福大学读心理学并获心理学博士学位,退休前是美国杜兰大学的心理学教授,退休后迁到加州居住,当时热衷于从事日本人学习英语的语言培训工作。由于本书涉及一些古英语,很难翻译,每当我与作者难以沟通时,泽千鹤子都能认真地帮我拆分句子,讲解结构,直到我能理解原文的意思为止。

本书第八章的内容是作者在与许多人合作研究的基础上写成的,在作者当时的工作秘书珍妮·迪尔多夫(Jeanne Deardorff)的帮助下,我与其中的一位合作者取得了联系,他叫吕忠林,是美籍华人,现在为美国南加州大学威廉·凯克(William M. Keck)认知神经科学讲座主任,心理学和生物医学工程教授,达纳(Dana)和戴维·多恩赛夫(David Dornsife)认知神经科学成像中心主任。他在南加州大学组建了脑处理实验室,主要研究目标是建立感知与认知的数学模型。吕忠林的研究领域主要包括:(1) 视觉与听觉认知、选择性注意、知觉学习的计算和心理物理研究;(2) 知觉、选择性注意、语言与决策的脑功能成像研究;(3) 阅读障碍、弱视与老年痴呆的视觉损伤。吕忠林于2002年被选为中国科学院海外评审专家,于2003年获得美国实验心理学会杰出青年奖,并成为该学会最年轻的高级会员。我们至今未曾谋面,但通过邮件,他为我解决过一些技术性术语的翻译问题。

本书最终能够在上海译文出版社出版,首先要真诚地感谢"全国社科规划办"项目管理政策的灵活性和时任上海译文出版

社副总编辑的赵月瑟老师的慧眼识珠。我与赵老师是于 2009 年 4 月 25 日在社科院哲学所的学术报告会上认识的。这个报告会由我主持，报告者是德国“诺贝尔替代奖”得主、慕尼黑大学物理学教授汉斯・彼得・迪尔（Hans-Peter Dürre），他受俞宣孟先生的邀请访问哲学所，学术报告的主题是：“科学与精神性——我们经验到的东西大于我们能掌握的东西”。在报告结束后，我有机会向赵老师介绍了本书的翻译工作，并打听是否有可能列入上海译文出版社的“二十世纪西方哲学译丛”出版的情况。几天后，我就幸运地接到了赵老师的邮件，要我把书的相关资料发给她。很快，本书就通过了出版社的审批。

在出版合同还没有签订之前，我第一次看到了“2009 年国家社科基金后期资助项目申报公告”，当时，本着试试看的态度，递交了本书的申报材料，并请童世骏教授、周昌忠教授和郦全民教授写了推荐信。之后，我便利用暑假到斯坦福从事合作研究，并在 8 月底回国时，顺便带回了美方授权上海译文出版社出版本书中译本的合同，并与上海译文出版社正式签订了出版合同。不久之后，在 2009 年的国家社科基金后期资助项目（批准号：09FZX009）的公告结果中，本书榜上有名。我看到这一消息，心中自然备感欣慰。可当我接到立项通知书时，新问题产生了，资助条件要求，译著应在规划办指定的出版社出版。如何作出选择使我为难。幸好，这种情况当时并非我一人。在我决定放弃后期资助时得知，社科规划办允许我们与出版社协商，只要出版社同意根据社科规划办的要求出版此书，就既不用放弃项目资助，也不用违约。在出版行业竞争激烈的今天，这无疑是一件两全其美的好事。在此，我再次特别感谢“全国社科规划办”的人性化管理与政策的灵活性。

本书从开始翻译到最终定稿大约有两年半时间，期间，责任

编辑王巧贞女士对本书作了认真细致的校译工作，并提出了许多好的建议，使本书的不当之处降到了最低。在书即将付梓之际，发自内心的感谢之言是必不可少的。我首先要感谢本书作者苏佩斯教授在我翻译过程中给予的很多帮助与支持；感谢他的秘书珍妮·迪尔多夫女士对我两次访问斯坦福大学所做的许多准备工作及其访问期间给予的工作上的各种帮助；感谢王瑞博士、泽千鹤子教授、吕忠林教授和德拉勒女士对我的翻译工作的帮助；感谢童世骏教授、周昌忠教授和郦全民教授的推荐信；感谢俞宣孟研究员、现任所长童世骏教授和副所长何锡蓉教授及其同事对我工作的支持；感谢上海译文出版社的赵月瑟副总编辑和社科编辑室主任马胜先生对出版本书的支持；感谢我的家人为我创造的和谐而宽松的家庭氛围。

翻译总是一件伴有遗憾的事情，虽然尽责尽力，但却难以尽善尽美。诚请专家读者对本书翻译的不当之处给予批评指正。

写于2010年11月5日上海松江书屋